



A University of Sussex PhD thesis

Available online via Sussex Research Online:

<http://sro.sussex.ac.uk/>

This thesis is protected by copyright which belongs to the author.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Please visit Sussex Research Online for more information and further details



DOCTORAL THESIS

The effect of dust in observing galaxies in the early Universe

*A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy*

in the

Astronomy Centre
School of Mathematical and Physical Sciences

Author:

Aswin P. Vijayan

Supervisor:

Peter A. Thomas & Stephen M. Wilkins

1st October 2021

असतो मा सद्गमया । तमसो मा ज्योतिर्गमया । मृत्योर्मा अमृतं गमया ।

From ignorance, lead me to truth. From darkness, lead me to light. From death, lead me to immortality.

– Brihadaranyaka Upanishad

ACKNOWLEDGEMENTS

As Tony Stark said “Part of the journey is the end”, and it has been a very enjoyable journey, because I got to share it with some incredible people. Looking back through all the archival memory data, there are a lot of people I like to thank for helping me reach here. I would like to start by thanking my supervisors Peter and Steve, who have been incredibly supportive and patient with my queries. I have learnt so much from you throughout these past 4 years. I have always enjoyed the countless chats we have had, research and otherwise. Mark Sargent, my progression reviewer, who also helped me in structuring my research. I would also like to thank Bruno Henriques, Chris Lovell, David Barnes, Dimitrios Irodou, James Trayford, Jussi Kuusisto, Maarten Baes, Peter Camps, Rob Yates, Scott Clay, Will Roper for their feedback and help during my research. Thanks also to all the people I worked with during my placement with ASTROCAST. Also a big thanks to Seb Oliver in his role in setting up DISCnet that funded my PhD as well as his help with my placement, and to the University of Sussex and the Science and Technologies Facilities Council (STFC). Finally, thanks to my examiners Romeel Davè and Seb Oliver for providing an enjoyable viva experience as well as for the constructive feedback.

This would be incomplete without mentioning all the people at the Astronomy Centre that have influenced my work in some way or the other. All the amazing people at the centre, whom I have had the opportunity to share company with, to name a few (I am gonna go alphabetically): Andrew, Alex, Azizah, Benoit, Carlos, Ciaran, Chris, Dan, Daniel, David, Dimitrios, 2×Edward, Ian, Itzi, Jussi, Kareem, Lucas, Luke, Maria, Michele, Pippa, Reese, Ridwan, Rose, Steven, Stephen, Sunayana, Will, and rest of the astro corridor. There was no shortage of fun on Falmer Friday’s, and the squash and badminton sessions. Also, thanks to my DISCnet friends, we had fun times at Old Thorns. Special thanks to my housemates during the Covid-19 lockdown, David and Lucas, help me keep sane during those insane times; the Rum Diaries including all the food and cocktails we tried, and the groovy times will always stay with me. Thanks also to my undergrad gang, ‘Jimmanz room’, for staying in touch even when everyone was in different parts of the globe.

I can’t thank my parents enough for supporting me through my aspiration to do astronomy, both emotionally and financially, without whose help, guidance and understanding, I would not be here today. Thanks also to my brother, Visakh and sister-in-law, Namitha. My favourite niece, Vaiga, made me smile everytime. Finally, I wish to thank my partner Mizuki, for supporting me through my PhD life. I loved all the adventures we shared and the random places we visited. Thank you for being so understanding and withstanding the difficulties of a long distance relationship during Covid-19 times.

Special thanks to Deliveroo, Just Eat and Uber Eats, and also to Disney+, Netflix and Prime Video.

DECLARATION

I, Aswin P. Vijayan, hereby declare that this thesis has not been and will not be, submitted in whole or in part to another university for the award of any other degree. The following parts of this submission have been published previously¹:

- Chapter 2: ‘Detailed dust modelling in the L-Galaxies Semi-Analytical Model of galaxy formation’, this has been published as [Vijayan et al. \(2019\)](#) in the Monthly Notices of the Royal Astronomical Society, Volume 489, Issue 3, November 2019, Pages 4072–4089.

Contributions: APV undertook the vast majority of the data analysis and produced the figures; APV and PAT worked on developing the grain growth framework building on the model provided by BMBH; SJC created the dust yield tables required for stellar dust production with the help of RMY, integrated the dust production and destruction framework into L-GALAXIES, did the initial data analysis and produced the first draft of the paper. PAT & RMY helped APV with the interpretation of the results and structuring the paper. PAT & SMW jointly supervised APV & SJC, and initiated the project.

- Chapter 3: ‘Introduction to First Light And Reionisation Epoch Simulations (FLARES)’, a major portion of this has been published as [Lovell et al. \(2021a\)](#) in the Monthly Notices of the Royal Astronomical Society, Volume 500, Issue 2, January 2021, Pages 2127–2145. I was the second author on this work and contributed significantly to the work.

Contributions: CCL carried out the majority of the writing for this work. APV, CCL, DI & WJR ran the re-simulations used in the study. APV wrote the code for selecting the regions, the weighting method to combine the regions, extracting relevant galaxy/particle properties to a single master file, merging of spurious objects to parent galaxies. PAT conceptualised the weighting scheme described in the paper. APV also helped in the editing of the paper narrative as well as with creation of figures and tables. PAT & SMW supervised CCL & APV, and initiated the project.

¹ Author short names – APV: Aswin P. Vijayan, BMBH: Bruno M. B. Henriques, CCL: Christopher C. Lovell, DI: Dimitrios Irodoutou, JK: Jussi Kuusisto, JT: James Trayford, MB: Maarten Baes, PAT: Peter A. Thomas, PC: Peter Camps, RMY: Robert M. Yates, SJC: Scott J. Clay, SMW: Stephen M. Wilkins, WJR: William J. Roper

- Chapter 4: ‘FLARES II: The photometric properties of high-redshift galaxies’, this has been published as [Vijayan et al. \(2021\)](#) in the Monthly Notices of the Royal Astronomical Society, Volume 501, Issue 3, March 2021, Pages 3289–3308.

Contributions: APV undertook the vast majority of the modelling, data analysis and produced the figures for this work. SMW wrote the module for the SED generation and helped APV with the use of it. CCL, PAT and SMW helped APV with the interpretation of the results, editing and structuring of the paper. PAT & SMW jointly supervised APV and initiated the project.

The following chapter is currently under peer-review:

- Chapter 5: ‘FLARES III: The properties of massive dusty galaxies at cosmic dawn’ has been submitted to MNRAS.

Contributions: APV undertook the vast majority of the modelling, running the SKIRT radiative transfer code, data analysis and produced the figures for this work. Some of the techniques used in the SKIRT modelling were implemented in earlier works by PC, JT and MB. CCL, JK, JT, MB, PAT, PC and SMW helped with editing the paper. PAT & SMW jointly supervised APV.

Brighton,

25th November 2021

Aswin P. Vijayan

Astronomy Centre
School of Mathematical and Physical Sciences
University of Sussex

DOCTORAL THESIS

The effect of dust in observing galaxies in the early Universe

by Aswin P. Vijayan

ABSTRACT

The first stellar populations kick-started the process of reionisation and began to enrich the pristine interstellar medium (ISM), with the first supernovae starting the dust formation and destruction processes. The pristine ISM evolved significantly over the course of the next billion years, affecting the subsequent evolution of the galaxies. Understanding this phase in the early Universe will help us learn how the galaxies evolved into the demographics we see today. Even in this regime, dust is an essential ingredient: even though the average dust content of galaxies is very low compared to the local Universe, it still has a significant impact on deriving meaningful answers from observations.

In this thesis I use a variety of numerical methods which include the semi-analytical methods (SAMs) and hydrodynamical simulations to study the evolution of dust in galaxies as well as its effect on the galaxy spectral energy distributions (SEDs). In the first section I use the L-GALAXIES SAM to incorporate a self-consistent model of dust formation and evolution. A novel feature in this work compared to similar efforts that have been published for semi-analytic and hydrodynamic models are (i) the more accurate consideration of the impact of molecular cloud chemistry on grain growth in dense molecular clouds (by separate tracking of dust in molecular and diffuse gas) and (ii) incorporating information on dust depletion fractions. I present the results of our implementation and compare it to the observational space.

In the second section I introduce FLARES (First Light And Reionisation Epoch Simulations), a suite of zoom simulations targeting a range of overdensities in the Epoch of Reionisation (EoR). The various overdensities were picked from a $(3.2 \text{ cGpc})^3$ volume, giving us access to some of the large scale modes in the Universe, which are and will be probed by current and next generation surveys/telescopes. These region were re-simulated using the EAGLE simulation physics, a well tested model in the low-redshift Universe. FLARES matches the stellar mass function and the star formation rate function of the current observations well. In the third section I show how we implement a simple line-of-sight (LOS) dust extinction model to retrieve the UV to near infrared SED including nebular emission from the FLARE simulations in the EoR. I present the UV luminosity function, the UV continuum slope (β) relations, the UV attenuation as well as the line luminosity and equivalent widths of some prominent nebular emission lines. The relative contribution of obscured and unobscured star formation is also explored, finding comparable contributions by $z \sim 6$. In the fourth section, I post-process the massive galaxies ($\geq 10^9 M_\odot$) in FLARES using the SKIRT radiative transfer code to study their dust properties such as the infrared luminosity function, the infrared excess - β (IRX- β) relation, various measures of luminosity-weighted dust temperatures. The FLARES IRX- β relation predominantly follows the local starburst relation. The luminosity-weighted dust temperatures increase towards higher redshifts, with the slope of the peak dust temperature - redshift relation showing a higher slope than the lower redshift relations obtained from previous observational and theoretical works.

CONTENTS

List of Figures	xiii
List of Tables	xvii
1 Introduction	1
1.1 Structure Formation	1
1.1.1 First Stars and Galaxies	2
1.1.2 Reionisation	3
1.2 Cosmological Simulations	5
1.2.1 N-body simulations	5
1.2.2 Semi-Analytical Modelling	6
1.2.2.1 L-Galaxies SAM	6
1.2.3 Hydrodynamical Simulations	9
1.2.3.1 The Eagle Simulations	11
1.3 Dust in Galaxies	14
1.4 Observations of the High Redshift Universe	16
1.4.1 Spectral Energy Distribution Modelling	17
1.4.1.1 Initial Mass Function	17
1.4.1.2 Stellar Population Synthesis	18
1.4.1.3 Nebular Emission	19
1.4.1.4 Fine-structure transitions	20
1.4.1.5 Dust Attenuation	21
1.4.1.6 SKIRT Radiative Transfer code	23
1.5 Thesis Overview	25

2	Detailed dust modelling in the L-GALAXIES semi-analytic model of galaxy formation	26
2.1	Introduction	26
2.1.1	Dust production and destruction	26
2.1.2	Previous modelling	27
2.1.3	Observational summary	28
2.1.4	Structure of the paper	29
2.2	The Model	30
2.2.1	Detailed chemical enrichment	30
2.2.2	Molecular gas	31
2.3	Detailed dust model	32
2.3.1	Supernova and AGB dust yields	33
2.3.2	Grain growth in molecular clouds	34
2.3.3	Dust destruction	37
2.3.4	Dust transfer	37
2.3.4.1	Star formation	37
2.3.4.2	Mergers	38
2.3.4.3	Other dust destruction mechanisms	38
2.4	Results: Dust Growth	38
2.4.1	Dust-to-Metal (DTM) ratio	39
2.4.1.1	DTM versus stellar mass	39
2.4.1.2	DTM versus metallicity	42
2.4.1.3	DTM fitting function	43
2.4.2	Integrated dust production rates	44
2.5	Results: Dust content of galaxies	47
2.5.1	Dust-to-Gas (DTG) ratio	48
2.5.2	Dust versus stellar mass	48
2.5.3	Dust mass function	51
2.6	Conclusions	53
2.A	Accretion Timescale	55
3	Introduction to First Light And Reionisation Epoch Simulations (FLARES)	58
3.1	Introduction	59
3.2	The FLARE Simulations	64
3.2.1	Region Selection	65

3.2.2	The Resimulation Method	66
3.2.3	Distribution Function Weighting	68
3.3	Results	69
3.3.1	Galaxy Number Counts	69
3.3.2	The Galaxy Stellar Mass Function	70
3.3.2.1	The cosmic GSMF	71
3.3.2.2	Environmental dependence of the GSMF	74
3.3.3	The Star Formation Rate Distribution Function	76
3.3.3.1	The cosmic SFRF	76
3.3.3.2	Environmental dependence of the SFRF	79
3.4	Conclusions	80
3.A	Selected regions	81
3.B	Fitted distribution functions	81
4	FLARES II: The Photometric Properties of High-Redshift Galaxies	84
4.1	The FLARE Simulations	84
4.1.1	Galaxy Identification	85
4.1.2	Metal Content	85
4.1.3	Spectral Energy Distribution Modelling	86
4.1.3.1	Stellar Emission	87
4.1.3.2	Nebular Emission	87
4.1.4	Dust Attenuation	87
4.2	Photometric Properties	90
4.2.1	UV Luminosity Function	90
4.2.1.1	LF creation	91
4.2.1.2	Luminosity Functions	91
4.2.1.3	Comparison with Observations and Models	95
4.2.2	UV continuum slope (β)	97
4.2.3	Effect of environment	101
4.2.4	Line Luminosities and Equivalent Widths	101
4.3	SFR distribution functions	104
4.4	Conclusions	106
4.A	Calibrating Dust Attenuation	107
4.B	UV LF	109
4.C	Other extinction curves	112

5	FLARES III: The properties of massive dusty galaxies at cosmic dawn	114
5.1	Introduction	115
5.2	Methods	117
5.2.1	Galaxy Identification and Selection	117
5.2.2	Spectral Energy Distribution modelling	118
5.3	Results	121
5.3.1	Luminosity functions	122
5.3.1.1	Environmental Dependence of IR LF	126
5.3.2	IRX- β	127
5.3.3	Dust temperatures	131
5.3.3.1	T_{peak}	132
5.3.3.2	T_{SED}	134
5.3.3.3	Redshift evolution of dust temperatures	135
5.4	Conclusions	138
5.A	Convergence Tests	140
5.B	Effect of AGN on the dust SED	141
5.C	Excluding birth cloud emission	143
5.D	T_{peak} and galaxy main-sequence	143
5.E	Comparison with line-of-sight model	144
6	Conclusions	146
6.1	Future Works	148
7	Acknowledgments	150
	Bibliography	152

LIST OF FIGURES

1.1	CMB map from the <i>Planck</i> satellite	2
1.2	Progress of reionisation in the Universe	4
1.3	Examples of some of the state of the art dark matter only and hydrodynamical simulations	10
1.4	Ingredients required to create galaxy SEDs	22
2.1	Evolution of the grain growth condensation fractions in molecular clouds and dif-fused medium	34
2.2	Carbon and oxygen depletion fractions for $z = 0$	35
2.3	The dust-to-metal ratio as a function of stellar mass from $z = 0 - 8$	39
2.4	The distribution of DTM ratios for $z = 0 - 8$	40
2.5	The dust-to-metal ratio as a function of stellar mass from $z = 0 - 8$, with galaxies coloured according to their stellar age	41
2.6	The dust-to-metal ratio as a function of their metallicity from $z = 0 - 8$	42
2.7	DTM fitting function in Equation 2.15 plotted against the DTM ratio from the model	43
2.8	Dust production and destruction rates as a function of redshift from the Millennium run	45
2.9	Variation of the dust production and destruction rate with stellar mass for $z = 0$	45
2.10	The dust-to-gas ratio as a function of stellar mass for $z = 0$	47
2.11	The dust-to-gas ratio as a function of gas phase metallicity for $z = 0$	47
2.12	The stellar-mass-dust-mass relation for redshifts $z = 0 - 8$	49
2.13	Dust Mass Function (DMF) for redshifts $z = 0$ and 1	52
2.14	Accretion timescale-stellar mass relation for redshifts $z = 0 - 8$	56

2.15	The median dust-stellar relation in our model for different values of $\tau_{\text{acc},0}$ at $z = 0, 2, 5$ and 6	57
2.16	DTM fitting function in Equation 2.15 is plotted against the DTM ratio for different $\tau_{\text{acc},0}$ values	57
3.1	Dark matter element resolution versus simulated volume	62
3.2	3.2 cGpc box from which we select our regions, compared to EAGLE and BLUEFIELDS volume	64
3.3	Visualisation of the dark matter integrated density in a number of resimulation regions of differing overdensity	65
3.4	Figures showing the probability distribution function of sampled overdensities and the distribution of overdensities within each simulation volume	67
3.5	Cumulative distribution of stellar masses for all FLARES regions combined and the fiducial EAGLE Reference volume	70
3.6	Redshift evolution of the galaxy stellar mass function of FLARES and EAGLE galaxies respectively	72
3.7	Parameter evolution for double-Schechter function fits to the FLARES composite galaxy stellar mass function and star formation rate function	73
3.8	FLARES composite galaxy stellar mass function evolution, alongside observational constraints and predictions from theoretical models	74
3.9	FLARES GSMF for $z = 5, 7$ and 9 split by overdensity bins. $z = 5$ GSMF for each individual region is also shown	75
3.10	Redshift evolution of the FLARES composite star formation rate distribution function	76
3.11	Evolution of the FLARES composite star formation rate distribution function, compared with observational constraints from UV data and other model predictions	78
3.12	FLARES SFRF for $z = 5, 7$ and 9 split by overdensity bins	79
4.1	FLARES composite galaxy stellar mass function compared with the 100 cMpc EAGLE Reference box for $z \in [5, 10]$	85
4.2	Mass weighted metallicities of the gas and stars of the FLARES galaxies at $z \in [5, 10]$ compared to observational constraints on the gas-phase metallicity	86
4.3	Line of sight tracing of the SPH density field	88
4.4	FLARES composite intrinsic and dust attenuated UV LF for galaxies in $z \in [5, 10]$	91
4.5	Schechter and double power-law fits to the FLARES UV LF	93

4.6	Redshift evolution of the parameters of Schechter and double power-law fits to the FLARES UV LF compared to values from observational and other theoretical studies	94
4.7	UV LF of the FLARES galaxies for $z \in [5, 10]$ compared to data from observational and theoretical studies from literature	96
4.8	UV continuum slope, β , plotted against the UV magnitude for $z \in [5, 10]$	97
4.9	The attenuation in the FUV plotted against the observed UV magnitude for $z \in [5, 10]$	99
4.10	The attenuation in the FUV plotted against the galaxy stellar mass for $z \in [5, 10]$	99
4.11	FLARES UV LF for $z \in [5, 10]$ split by binned log-overdensity	100
4.12	Predictions for the luminosity function and equivalent widths of 6 prominent UV and optical lines in FLARES for $z \in [5, 10]$	102
4.13	Predicted distribution of combined $H\beta$ and $[\text{OIII}]\lambda_{4959,5007}$ equivalent widths and line luminosities against the stellar masses and UV luminosities for FLARES galaxies at $z \sim 7, 8$	102
4.14	The De Barros et al. (2019) and predicted combined $H\beta$ and $[\text{OIII}]\lambda_{4959,5007}$ line luminosity function of FLARES galaxies at $z \sim 8$.	103
4.15	Predicted $[\text{CIII}]\lambda_{1907, \lambda_{1909}}$ line equivalent widths of FLARES galaxies at $z \sim 7$	103
4.16	FLARES composite galaxy total, obscured and unobscured star formation rate function for $z \in [5, 10]$	105
4.17	FLARES composite galaxy total, obscured and unobscured star formation rate density for $z \in [5, 10]$	105
4.18	UV continuum slope β for different values of κ_{BC} at $z = 5$	108
4.19	Same as Figure 4.13, now showing the line luminosity and equivalent width for different values of κ_{BC}	109
4.20	Same as Figure 4.9 and 4.10, now showing the attenuation as a function of intrinsic UV luminosity.	109
4.21	β values and attenuation in the far-UV respectively for different extinction curves at $z = 5$	112
4.22	Same as Figure 4.13, now showing the line luminosity and equivalent widths for different extinction curves.	113
5.1	Relationship between the galaxy stellar mass and the star formation rate of the galaxy sample for $z \in [5, 10]$	118
5.2	Evolution of the median dust-to-metal ratio with the galaxy stellar mass across $z \in [5, 10]$	119
5.3	UV luminosity function of the FLARES galaxies for $z \in [5, 10]$ obtained using SKIRT	122

5.4	IR luminosity function of the FLARES galaxies for $z \in [5, 10]$	123
5.5	Rest-frame 250 μ m luminosity function of the FLARES galaxies for $z \in [5, 7]$	125
5.6	FLARES IR LF for $z \in [5, 7]$ split by binned log-overdensity	126
5.7	IRX- β distribution of the FLARES galaxies for $z \in [5, 10]$	128
5.8	IRX-stellar mass distribution of the FLARES galaxies for $z \in [5, 10]$	130
5.9	Variation of λ_{peak} (T_{peak}) with the galaxy stellar mass, total infrared luminosity and the specific star formation rate respectively for $z \in [5, 10]$	132
5.10	Evolution of the peak dust temperature (T_{peak}) and the SED dust temperature with redshift from the FLARE simulation	136
5.11	Shows the SED of a galaxy at $z = 8$ for a few different configurations of the SKIRT code	140
5.12	Shows the SED of a galaxy at $z = 8$ for the SMC and Milky Way grain distribution.	140
5.13	Shows how the total infrared luminosity function changes, if all the energy from the SMBH accretion went into the infrared	142
5.14	UV and IR luminosity functions with and without the birth cloud attenuation for $z = 8$	142
5.15	Variation of λ_{peak} (T_{peak}) with the IR luminosity for $z \in [5, 9]$, with the FLARES galaxies split based on their SFR	143
5.16	UV LF obtained from the LOS dust extinction model compared to the SKIRT implementation	145
6.1	Spread in the UV attenuation for different lines-of-sight as a function of the dust surface density for $z \in [5, 8]$	149

LIST OF TABLES

2.1	The conversion efficiencies used for the production of dust grains in supernovae remnants based on the mass return of key metals	34
3.1	Regions selected from the parent volume for resimulation. We provide their positions within the parent volume, their overdensity δ as defined by Equation 3.1, their <i>rms</i> overdensity σ , and weights, f_j , calculated as per Section 3.2.3.	82
3.2	Best fitting double-Schechter function parameter values for the Galaxy Stellar Mass Function. α_2 is fixed at -1	83
3.3	Best fitting double-Schechter function parameter values for the Star Formation Rate distribution function. α_2 is fixed at -1	83
4.1	Binned UV LF values for the FLARES galaxies	110
4.2	Best-fitting Schechter and double power-law function parameter values for the observed UV LF	111

INTRODUCTION

1

This thesis addresses the question on how dust influences the various properties of the galaxies we observe in the early Universe. The different chapters are focussed on developing a model for dust in simulations of galaxy formation and evolution, in semi-analytical models and hydrodynamical simulations. The main theme connecting all the chapters is the growth of dust in galaxies through the Universe and how this influences the observation of galaxies in the early Universe. In this chapter I will provide a brief review of structure formation in the Universe, simulating these structures as well as modelling the emission from galaxies and the complexity dust brings to the picture.

1.1 Structure Formation

The current standard cosmological model, is known as the Lambda Cold Dark Matter (Λ CDM) model, with its energy components being baryonic matter, cold dark matter and dark energy, with dark energy responsible for the accelerated expansion of the Universe (Peebles & Ratra, 2003). In the standard model of cosmology, structures emerge in this expanding space-time due to small density perturbations in an otherwise homogeneous and isotropic Universe. These structures arose from initial quantum fluctuations that were stretched out to macroscopic scales during inflation (which began $\sim 10^{-36}$ s after the Big Bang) and then froze out once they exited the horizon (see Guth, 1981; Mukhanov & Chibisov, 1981; Linde, 1982; Peebles & Ratra, 2003). The process of Big Bang Nucleosynthesis (BBN) created the primordial elements in the Universe a few minutes after the Big Bang (see Wagoner et al., 1967; Steigman, 2007). The Universe was still in a state of hot plasma at this stage and radiation dominated.

As the Universe expanded and cooled down, baryonic matter eventually became decoupled

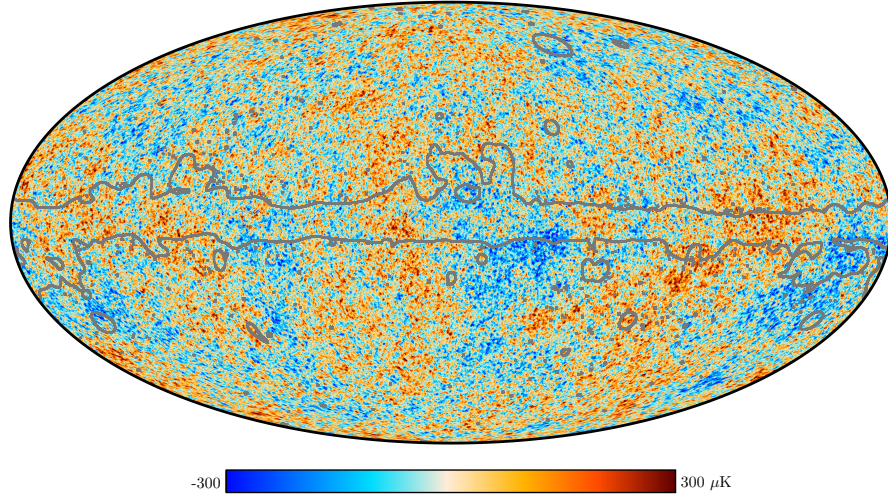


Figure 1.1: The CMB map from the *Planck* satellite with the scale of the temperature inhomogeneities indicated at the bottom. Also shown is the galactic plane in grey lines. Courtesy: <https://www.cosmos.esa.int/web/planck/picture-gallery>.

from radiation (redshift, $z \sim 1100$). This background radiation travelled (mostly) unimpeded through the Universe, and is observed today in the microwave frequency, due to the subsequent expansion of the Universe, as the Cosmic Microwave Background (CMB) radiation. Analysis of the CMB provides very stringent constraints on the Λ CDM parameters or any other cosmological model.

Dark matter ultimately drives the formation of structure in the Universe. In the Λ CDM model of structure formation, dark matter is attracted to local peaks of initial density fluctuations (very small anisotropies seen in the CMB, see Figure 1.1) that were formed during inflation. This happens during the matter domination era, creating regions of higher dark matter density. These are seeds to the formation of the first systems in gravitational equilibrium, the extended structure of dark matter called halos. They form in an hierarchic way, with the smaller halos collapsing first, which then aggregate to form the bigger systems, referred to as hierarchical clustering. When baryonic matter eventually decouples from radiation, it collapses into these dark matter overdensities.

It is also worth noting that while Λ CDM is the currently preferred model, alternatives exist, for example those (e. g. MOND, [Milgrom, 1983](#)) that replace the need for dark matter and energy by a modification to gravity.

1.1.1 First Stars and Galaxies

These first systems provide hosts to the first stars and galaxies formed in the Universe. The baryons which are present inside these dark matter halos can cool through radiative processes and

start to sink to the centre. Star formation is expected to take place in clouds of dense molecular hydrogen. This happens when the cloud collapses under its own gravity, fragments and then form stars when the Jeans mass is reached.

The first stars were thought to form in minihalos (mass of $\sim 10^6 M_\odot$) at $z \sim 20-30$, when the age of the Universe was a few hundred million years (see [Bromm & Larson, 2004](#); [Yoshida et al., 2012](#), for detailed reviews). These first stars were born at the heart of the first galaxies. These first generation stars, so called the Population III stars, were born from pristine or primordial, metal free gas cooled from the hydrogen molecule lines. As a result of this, their Jeans mass is expected to be significantly higher than the molecular clouds in the local Universe, thus making these generation of stars likely to be dominated by massive, short lived objects, whose lifecycle ended in massive explosions ([Bromm & Larson, 2004](#)). They kickstarted the metal and dust enrichment of the early Universe, thus creating more cooling pathways from metal line cooling. This paved way for the currently seen crop of stars referred to as Population II or I based on their metal content ([Bromm, 2013](#); [Klessen, 2019](#)).

The formation of the first stars is also expected to be in tandem with the formation of the first black holes, which can be from stellar remnants and/or direct collapse black holes from metal free gas (see [Bromm & Loeb, 2003](#); [Begelman et al., 2006](#)). Direct collapse black holes could be the source of the supermassive black holes (SMBH) at the centre of galaxies seen in the early Universe ($z > 6$). SMBH with high enough accretion, called Active Galactic Nucleus (AGN), can heat the surrounding gas, preventing it from forming stars. Thus they can play an active part in regulating the growth of massive galaxies.

1.1.2 Reionisation

Reionisation is the last phase-transition in the Universe, with neutral hydrogen transitioning to being mostly ionised. This phase-transition is the direct effect of the formation of first stars and galaxies which heat and ionise their surroundings and subsequently the intergalactic medium (IGM). The exact details on the progress of reionisation is unclear, for example when it exactly started and ended, and which sources (stars or Active Galactic Nuclei as well as the mass of galaxies responsible) were the main drivers and the topology of these ionised regions. Some of the current constraints suggests that the process takes place over $\sim 1\text{Gyr}$ (from $z \sim 15 \rightarrow z \sim 5$) in time ([Zaroubi, 2013](#)), also see Figure 1.2.

The first stars create their own HII regions within the galaxies slowly ionising the ISM. These galaxies then start to create an ionised bubble of their own which are separated by vast neutral

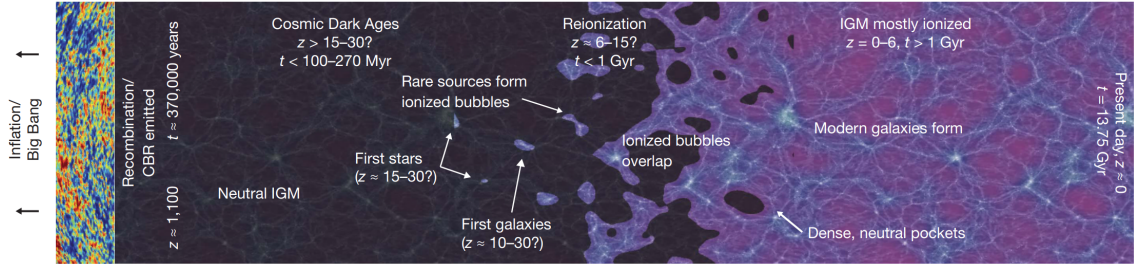


Figure 1.2: Transition of the Universe from fully neutral to to mostly ionised. After recombination, the Universe enters the cosmic dark ages, with no radiation source except for the cosmic background radiation, currently observed in the microwave. The first galaxies and stars started creating local ionised bubbles. As their abundance increased, these bubbles increasingly overlapped and more regions became ionised, and subsequently completing the process of reionisation. Reproduced from [Robertson et al. \(2010\)](#).

IGM. These bubbles start to grow with escaping ionising photons from the galaxy that can also include quasar sources. As the bubbles start to grow, they overlap with other regions ([Meiksin & Madau, 1993](#); [Gnedin, 2000](#)). Several galaxies contribute to the ionisation, accelerating the process of reionisation, until most of the IGM is ionised and only few neutral patches remain. With the propagation of the ultraviolet (UV) background into these pristine environments, reionisation is finally complete (figure 1.2).

As mentioned earlier, the sources driving reionisation are still unclear. The current leading theory is that it is the abundance of low-mass galaxies in the early Universe with higher escape fraction of ionising photons than the galaxies in the local Universe that is responsible. There are also other new avenues being probed such as the effect of binary stars which can substantially increase the number of escaping photons ([Stanway, 2017](#)) as well as changes to the initial mass function (IMF) in some of the massive galaxies. With the help of observations of the rest-frame UV using the *Hubble space telescope* (*HST*), there have been significant inroads made into understanding the shape of the UV luminosity function at high-redshift. Gravitationally lensed fields have helped to probe the fainter galaxies, providing more statistical power to the low-mass/faint end of the galaxy stellar mass/UV luminosity function. The jury is still out on the presence of a turn-over and where exactly this occurs in these functions at high-redshift, which will have an impact on the progress of reionisation. There can also be contribution to reionisation from AGN residing in the most massive galaxies. However, their contribution is assumed to be negligible owing to their low number densities at high-redshift.

1.2 Cosmological Simulations

Simulations have become an essential part of every discipline like astrophysics, particle physics, biology or instrumentation. These simulations have become in the last couple of years one of the most effective tool to study and solve astrophysical problems. Dark matter and dark energy which are the main energy components of the geometrically flat Universe are also ingredients of these simulations, even though we do not understand their exact nature. Simulations can make useful predictions just by knowing their general characteristics. The baryonic component that makes up $\sim 4\%$ of the energy content are crucial to model galaxies and other processes that occur at the small scales. By modelling these components, simulations help us in creating a window to the underlying physics that goes on at different scales, both in spatial and energy scales in the Universe. These when compared to observational data help us improve our models as well as understand the physical processes that shape galaxy formation and evolution.

The following are the types of frameworks commonly used to get an understanding of the observations: (i) N-body simulations that follow dark matter only as they form halos and other substructures, (ii) Semi Analytical models (SAM) which uses analytical prescriptions for baryonic physics usually built on top of some dark matter only simulation and (iii) Hydrodynamical simulations that follow both dark matter and baryonic matter. I will briefly describe these in the following sub-sections.

1.2.1 N-body simulations

In the Λ CDM model of the Universe, most of the matter is in the form cold dark matter ($\sim 85\%$). This is very advantageous when running dark matter only simulations since they can be treated as a collisionless fluid, interacting only through gravity. This makes the calculation involved in their interactions easier and faster to compute, thus allowing for simulating huge representative volumes at very high dark matter mass resolutions.

The N-body method traces the motion of each particle in the simulation numerically, by solving the mutual gravitational forces between them. For N number of particles, when N is large, instead of computing the gravitational field by summing over all the individual particle contributions, it is computationally more efficient to group particles to other particles according to their distances. In such a case the force the group exerts on a single particle is given by the first terms in the force multipole expansion. This method is referred to as the tree algorithm (Barnes & Hut, 1986). Another method is to treat the gravitational field by discretising it on a

cartesian grid or ‘mesh’ (Hockney & Eastwood, 1988). Forces are then computed by interpolating the derivatives of the potentials to the particle positions. Even though the particle-mesh (PM) method is faster than the tree algorithm as well as being able to account for periodic volumes by default, it is not well suited at small scales due to limited resolution (as it is limited in resolution to their cell size). So usually a hybrid Tree-PM algorithm is adopted where the short-range forces are computed using the tree algorithm while the long range forces with the PM method.

In this thesis the N-body dark matter only periodic volume simulations (Millennium (Springel et al., 2005a), Millennium II (Boylan-Kolchin et al., 2009), MACSIS dark matter volume (Barnes et al., 2017a)) that are being used were all run with the GADGET code (Springel et al., 2005a), which uses the Tree-PM method. Other than periodic volumes, high resolution dark matter only zoom simulations of smaller regions, for e. g. Milky Way sized halos (e. g. Via Lactea II (Diemand et al., 2008), Aquarius (Springel et al., 2008)) or the local group (e. g. Sibelius (Sawala et al., 2021)) have also been undertaken.

1.2.2 Semi-Analytical Modelling

One of the main ingredient in galaxy formation is baryons which cool through radiative processes to form stars and black holes. Semi-Analytical Models (SAMs) provide a computationally less expensive way than hydrodynamical simulations to self-consistently evolve the baryonic components associated with dark matter merger trees that are derived from numerical simulations or Press-Schechter merger calculations. SAMs use coupled differential equations to follow the baryonic physics like star formation, chemical enrichment, different feedback mechanisms, etc involved in galaxy formation and evolution within the dark matter halos (White & Frenk, 1991; Baugh, 2006). A number of SAMs such as L-GALAXIES (Kauffmann et al., 1999; Springel et al., 2001; De Lucia et al., 2004; Angulo & White, 2010; Guo et al., 2011; Henriques et al., 2015; Clay et al., 2015; Henriques et al., 2020, etc), GALFORM (Cole et al., 2000; Lacey et al., 2011; Gonzalez-Perez et al., 2014; Lacey et al., 2016), SAGE (Croton et al., 2006; Croton et al., 2016) have been developed over the years to study the various galaxy properties and scaling relations. In this thesis I use L-GALAXIES or the Munich SAM as described below.

1.2.2.1 L-Galaxies SAM

The Munich SAM or L-GALAXIES has been developed over the years to include the relevant processes required for galaxy evolution. These are a suite of semi-analytical models for galaxy formation implemented on the Millennium (Springel et al., 2005a) and the Millennium-II (Boylan-

Kolchin et al., 2009) simulations. The Millennium simulation traces the dark matter evolution in a cubic box of side length $500h^{-1}$ cMpc. This box is sensitive to dark matter halos of mass $\gtrsim 10^{10}h^{-1}M_{\odot}$ due to resolution limits because of its large size. In order to study lower mass halos the simulation was rerun with a smaller box of length $100h^{-1}$ cMpc with the same number of particle, thus having higher mass resolution, the Millennium-II simulation. The simulations assume the Λ CDM cosmology parameters derived from the combined analysis of WMAP1 and 2dFGRS. Both the simulations trace the interaction of 2160^3 particles from $z = 127$ to $z = 0$. For this work the cosmological parameters are as follows: $\sigma_8 = 0.829$, $H_0 = 67.3 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_{\Lambda} = 0.685$, $\Omega_m = 0.315$, $\Omega_b = 0.0487$ and $m = 0.96$. The cosmology was scaled following the Angulo & White (2010) technique, as updated by Angulo & Hilbert (2015), to represent the best-fit cosmological parameters derived from the year one Planck cosmology data (Planck Collaboration et al., 2014).

The version of the L-GALAXIES model used in this thesis, in Chapter 2 is based on the Henriques et al. (2015) version. The galaxy formation model includes physical prescriptions for processes such as gas cooling, star formation, supernova feedback, formation and growth of black holes, AGN feedback and galaxy interactions and mergers. The simulation uses the Chabrier initial mass function Chabrier (2003). Parameters (17 in total, see Table S1 in Henriques et al., 2015) such as for star formation efficiency, supernovae feedback, black hole growth and feedback, etc are constrained using a Markov Chain Monte Carlo (MCMC) approach against a set of observational constraints (stellar mass function and the passive galaxy fraction as a function of stellar mass and redshift, $z < 3$). Here, I describe very briefly some of the model characteristics:

- The model assigns a baryonic fraction (f_b^{cosmic}) of 15.5% to collapsed halos. To model photoheating by the UV background heating, a filtering mass ($M_F(z)$) is adopted for halos which have mass below this value to have their baryon fraction reduced with respect to the universal fraction. This mass is adopted from Gnedin (2000) and is modelled as follows

$$f_b(M_{\text{vir}}, z) = \frac{f_b^{\text{cosmic}}}{\left(1 + (2^{\alpha/3} - 1) \left[\frac{M_F(z)}{M_{200}}\right]^{\alpha}\right)^{3/\alpha}}, \quad (1.1)$$

where M_{200} is the virial mass of the halo and α is a free parameter, which in this case is adopted a value of 2. $M_F(z)$ is a function of redshift, varying from $\sim 6.5 \times 10^9 M_{\odot}$ at $z = 0$ to $\sim 10^7 M_{\odot}$ at $z = 8$ (Okamoto et al., 2008). Thus for large halos with $M_{200} \gg M_F(z)$, there is negligible suppression in the baryon fraction, while for smaller halos (in our case many of the resolved halos in Millennium-II) with $M_{200} \ll M_F(z)$, the baryon fraction is reduced.

- Following [White & Frenk \(1991\)](#); [Springel et al. \(2001\)](#), the gas is assumed to be in a quasi-static state, cooling from the hot halo where its distribution is isothermal. Gas cooling happens within a radius where gas cooling time equals the halo dynamical time.
- The angular momentum and hence the size of the stellar and gaseous disk are computed using the prescription in [Guo et al. \(2011\)](#). Their profile is assumed to be exponential with the evolution of both the components modelled separately. Bulges can form through minor and major mergers as well as disk instabilities.
- Stars are assumed to form from the cold gas within the galaxy disks, with the star formation rate given by

$$\dot{M}_\star = \alpha_{\text{SF}} \frac{M_{\text{gas}} - M_{\text{crit}}}{t_{\text{dyn,disk}}}, \quad (1.2)$$

where $\alpha_{\text{SF}} = 0.055$, M_{gas} is the total cold gas mass, $t_{\text{dyn,disk}} = R_\star/V_{\text{max}}$, is the dynamical time of the disk (R_\star is the stellar disk scale length and V_{max} is the maximum circular velocity of the halo). M_{crit} is a threshold mass for star formation (following [Kauffmann et al., 1996](#)) given by

$$M_{\text{crit}} = M_{\text{crit},0} \left(\frac{V_{200,c}}{200 \text{ km/s}} \right) \left(\frac{R_{\text{gas}}}{10 \text{ kpc}} \right), \quad (1.3)$$

where $M_{\text{crit},0} = 3.8 \times 10^9 M_\odot$, $V_{200,c}$ is the halo virial velocity and R_{gas} is the gas disk scale length. Mergers can also trigger starbursts, modelled following [Somerville et al. \(2001\)](#).

- The total mass of metals is tracked, with the yield i. e. mass of metals per solar mass a free parameter in the MCMC.
- Supernovae feedback in the model injects energy into the ISM, part of which is used to heat the cold gas and inject it into the hot gas reservoir. The remaining energy is used to eject material to an external reservoir which may or may not be reincorporated later (delayed reincorporation). See S1.7 in [Henriques et al. \(2015\)](#) for more details.
- Black hole (BH) growth is modelled following [Croton et al. \(2006\)](#), with the build up of mass happening through two channels. The first, *quasar* mode is when they grow through galaxy mergers, in the instance where both the galaxies hosts central BHs. This mode mostly drives the BH mass growth, but is not associated with any feedback. The second one termed *radio* mode where the accretion of gas from hot gas reservoirs, inject energy into the hot atmosphere driving hot bubbles and jets. This mode has negligible effect on the BH mass growth, but is responsible for efficient feedback that suppress star formation

above the knee in the stellar mass function. See S1.10 in [Henriques et al. \(2015\)](#) for more details.

In chapter 2, the detailed chemical enrichment introduced in [Yates et al. \(2013\)](#) for L-GALAXIES is used, which follows individual elements in the different gas phases as well as includes a framework for delayed enrichment of metals from supernovae and stellar winds. It should also be noted that the most recent version of L-GALAXIES ([Henriques et al., 2020](#)) resolves galaxies by splitting them in 12 concentric annuli of fixed radius and width, and accounts for the migration of material between them.

1.2.3 Hydrodynamical Simulations

Hydrodynamical simulations differ from SAMs, such that they self-consistently model the evolution of the baryonic component in addition to dark matter. This adds in more complexity to the code, due to the need to solve the equations of hydrodynamics concurrently with the gravity solver. These simulations allow for the exploration of the properties of the baryonic component spatially. They also have prescriptions for gas cooling, star formation, chemical enrichment from stars, feedback mechanisms, etc.

An important thing to keep in mind is that the particles/resolution elements in a simulation does not describe individual dark matter or baryonic particles. They actually have masses that are many orders of magnitude larger. Thus they can not resolve every physics that happen at the small scales and thus does not exactly mimic reality. Present day simulations involve ‘subgrid’ models to deal with some of these unresolved processes and couple them to resolved scales. These subgrid models are very similar to SAMs, also requiring fine-tuning of the involved free parameters. The implementation and parameterisation of subgrid routines is one of the greatest source of uncertainty in cosmological simulations, and adjustment of these characteristics can result in the dramatic alteration of simulation outcomes. So when interpreting simulation results it is always good to compare with observations or analytical models to test its reliability.

In the past decade a number of state of the art hydrodynamical simulations such as MASSIVE-BLACK ([Matteo et al., 2012](#)), ILLUSTRIS ([Vogelsberger et al., 2014a,b](#); [Genel et al., 2014](#); [Sijacki et al., 2015](#)), HORIZON-AGN ([Dubois et al., 2014](#)), MASSIVEBLACK-II ([Khandai et al., 2015](#)), EAGLE ([Schaye et al., 2015](#); [Crain et al., 2015](#)), BLUETIDES ([Feng et al., 2016](#)), MUFASA ([Davé et al., 2016](#)), COSMIC DAWN ([Ocvirk et al., 2016](#)), ILLUSTRIS-TNG ([Naiman et al., 2018](#); [Nelson et al., 2018](#); [Marinacci et al., 2018](#); [Springel et al., 2018](#); [Pillepich et al., 2018](#)), SIMBA ([Davé et al., 2019](#)), COSMIC DAWN II ([Ocvirk et al., 2020](#)), etc have been developed independently to study galaxy forma-

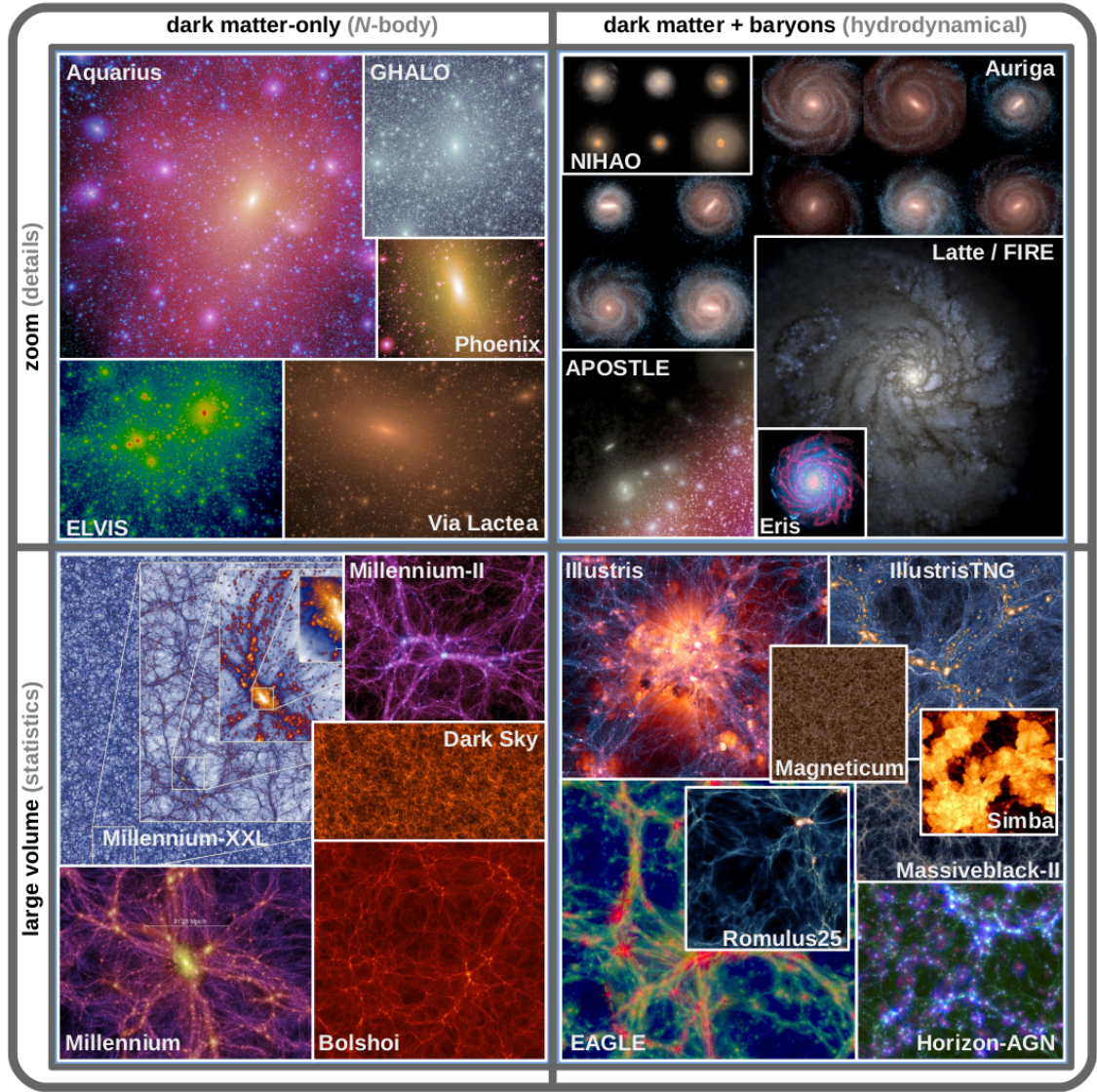


Figure 1.3: The figure shows a few examples of state of the art periodic and zoom, dark matter only as well as hydrodynamical simulations that have been undertaken. Reproduced from [Vogelsberger et al. \(2020a\)](#).

tion and evolution in a representative volume of the Universe. Other than these large volume cosmological simulations, there are also zoom simulations which focus on individual galaxies (e. g. AURIGA ([Grand et al., 2017](#)), VELA ([Ceverino et al., 2014](#)), etc) or halos/small volumes (e. g. APOSTLE ([Sawala et al., 2016](#)), FIRE ([Wetzel et al., 2016](#)), SPHINX ([Rosdahl et al., 2018](#)), etc) resolving the smaller scales in more detail. A concern with zoom simulations of single galaxies or small regions is that it is complicated (due to extremely high computational costs) to test if the employed physics would produce galaxy populations that are representative of the Universe. Figure 1.3 (reproduced from [Vogelsberger et al., 2020a](#)) shows some examples of dark matter only and hydrodynamical simulations that have been undertaken.

In this thesis I employ the EAGLE simulation physics to study the high-redshift Universe.

Below I briefly describe the EAGLE simulation scheme.

1.2.3.1 The Eagle Simulations

The EAGLE (Evolution and Assembly of GaLaxies and their Environment [Schaye et al., 2015](#); [Crain et al., 2015](#)) is a state of the art hydrodynamical simulation that uses the Smooth Particle Hydrodynamics (SPH, see reviews by [Monaghan, 1992](#); [Springel, 2010a](#), for details), the most popular Lagrangian method. The SPH technique solves the Euler equations (representing the conservation of mass, momentum and energy, usually closed by assuming a polytropic equation of state) by representing the fluid by a set of mass elements or particles and following their motion. In this representation, any continuous fluid quantity at a particular position, can be represented by a smoothed interpolated version ($F_s(\mathbf{r})$), obtained via kernel interpolation from neighbouring particles within a smoothing length h ,

$$F_s(\mathbf{r}) = \sum_j \frac{m_j}{\rho_j} F_j W(\mathbf{r} - \mathbf{r}_j, h). \quad (1.4)$$

Here, m is the particle mass, ρ its density, and W is the kernel, which is a spherical function of the distance between the particles in units of the smoothing length. One of the advantages of SPH is that it can track the movement of mass directly, easier to track inflows and outflows in galaxies. It also suffers from certain disadvantages, such as the difficulty to track shocks and mixing of phases (see reviews by [Springel, 2010a](#); [Somerville & Davé, 2015](#), for more details on successes and limitations of SPH).

It should also be noted that, cosmological simulations also use the traditional Eulerian methods, that solve equations on a discretised grid frame than the fluid frame. A very widely used implementation of such method is Adaptive Mesh Refinement (AMR), where cells satisfying some local criteria is split into subcells, enabling higher resolution in those regions, with the Riemann problem solved across the face of the cell. An implementation of this method for cosmological simulations was implemented in the RAMSES code ([Teyssier, 2002](#), used to run the HORIZON-AGN simulation). There are also hybrid codes, for e. g. AREPO ([Springel, 2010b](#), used to run the ILLUSTRIS and ILLUSTRIS-TNG simulations), that uses Voronoi tessellation to subdivide space around particles, with the mesh regenerated as the fluid moves. Another hybrid code is GIZMO ([Hopkins, 2015](#), used in running the MUFASA and SIMBA simulations), uses meshless finite mass and volume methods, and have been successfully applied to astrophysical simulation problems.

The EAGLE suite of simulations were run with a Planck year 1 cosmology with $\Omega_m = 0.307$, $\Omega_\Lambda = 0.693$, $\Omega_b = 0.04825$, $h = 0.6777$ and $\sigma_8 = 0.8288$. The initial conditions were generated using the second-order Lagrangian perturbation theory method of [Jenkins \(2010\)](#) and the

gaussian white-noise field generating code PANPHASIA (Jenkins, 2013). The simulations were run using a modified version of the N-body Tree-PM SPH code P-GADGET-3, which was last described in Springel et al. (2005a). The main modifications were done to the formulation of SPH, the time stepping and most importantly the subgrid physics. The simulations use the SPH scheme which is collectively termed as ANARCHY (Schaller et al., 2015). The models were run in boxes of length 25–100 h^{-1} cMpc and employ a resolution that is sufficient to marginally resolve the Jeans scale in the warm ($T \sim 10^4$ K) ISM. The fiducial Eagle simulation (100 h^{-1} cMpc a side), the recalibrated high-resolution (25 h^{-1} cMpc a side) and the higher AGN heating temperature model is referred to as Reference (Ref), Recalibrated (Recal) and AGNdT9 respectively. The simulations use the Chabrier initial mass function (Chabrier, 2003). The subgrid physics is based on that was developed for the OWLS (Schaye et al., 2010), which was also used in GIMIC (Crain et al., 2009) and COSMO-OWLS (Le Brun et al., 2014) simulations. The details of the model is described extensively in Schaye et al. (2015); Crain et al. (2015); Schaller et al. (2015), I describe very briefly some of them here:

- The model uses the \mathcal{C}_2 kernel described as

$$W(r, h) = \frac{21}{2\pi h^3} \begin{cases} (1 - \frac{r}{h})^4(1 + 4\frac{r}{h}) & \text{if } 0 \leq r \leq h \\ 0 & \text{if } r > h. \end{cases} \quad (1.5)$$

It adopts the time-step limiter implemented in Durier & Dalla Vecchia (2012), which ensures that sudden changes in the particle internal energy is captured and propagated to neighbouring particles. EAGLE also adopts the pressure-entropy SPH formalism of Hopkins (2013) to derive the equations of motion as well as an artificial viscosity (Cullen & Dehnen, 2010), and an artificial conductivity switch (e. g. Price, 2008).

- Element-by-element radiative cooling and photoheating implemented following Wiersma et al. (2009a), under the assumption that the gas is optically thin, is in ionization equilibrium, and is exposed to the CMB radiation and a spatially uniform, temporally-evolving Haardt & Madau (2001) UV/X-ray background. The UV/X-ray background is imposed instantaneously at $z = 11.5$ (from Planck Collaboration et al., 2014, constraints on reionisation).
- Pressure dependent star-formation recipe, implemented stochastically following Schaye & Dalla Vecchia (2008), with a metallicity dependent star formation threshold proposed in Schaye (2004). More details in section 4.3 of Schaye et al. (2015).

- Star particles are treated as simple stellar populations with the Chabrier IMF (Chabrier, 2003). Stellar evolution and mass loss is implemented based on Wiersma et al. (2009b).
- Feedback from stars is only thermal, which is implemented stochastically following Dalla Vecchia & Schaye (2012), where the temperature increment of the heated resolution element is specified ($\Delta T_{\text{SF}} = 10^{7.5} \text{M}_{\odot}$). The subgrid radiative losses are dependent on the local ISM conditions. More details in section 4.5 of Schaye et al. (2015).
- Black hole (BH) seeds are placed at the centre of every halo with total mass greater than $10^{10} \text{M}_{\odot}/h$ that does not already contain a BH (Springel et al., 2005b), by replacing the highest density gas particle, inheriting the particle mass and acquiring a subgrid BH mass, $m_{\text{BH}} = 10^5 \text{M}_{\odot}/h$. BH gas accretion is dependent on the mass of the BH, the local density and temperature, the velocity of the BH relative to the ambient gas, and the angular momentum of the gas with respect to the BH. Specifically, BH accretion rate is given by the minimum of the Eddington and Bondi-Hoyle accretion rates (Bondi & Hoyle, 1944) times an efficiency factor,

$$\dot{m}_{\text{BH}} = (1 - \epsilon_r) \min(\dot{m}_{\text{Edd}}, \dot{m}_{\text{Bondi}} \times \min(C_{\text{visc}}^{-1} (c_s/V_{\phi})^3, 1)), \quad (1.6)$$

where

$$\dot{m}_{\text{Edd}} = \frac{4\pi G m_{\text{BH}} m_p}{\epsilon_r \sigma_T c}, \quad (1.7)$$

$$\dot{m}_{\text{Bondi}} = \frac{4\pi G^2 m_{\text{BH}}^2 \rho}{(c_s^2 + v^2)^{3/2}}, \quad (1.8)$$

and m_p is the proton mass, σ_T is the Thomson cross section, c the speed of light, $\epsilon_r = 0.1$ is the radiative efficiency of the accretion disc, c_s is the sound speed, and v the relative velocity of the BH and the gas. V_{ϕ} is the rotation speed of the gas around the BH (see equation 16 in Rosas-Guevara et al., 2015) and C_{visc} , a free parameter related to the viscosity of the (subgrid) accretion disc which for Recal and AGNdT9 use a value 10^3 and 10^2 times higher than the Reference volume respectively. More details in section 4.6.1 and 4.6.2 of Schaye et al. (2015).

- Similarly to stellar feedback, AGN feedback is thermal and implemented stochastically, following Rosas-Guevara et al. (2015), resembling a quasar mode style feedback. The Reference EAGLE simulations adopt a heating temperature of $\Delta T_{\text{AGN}} = 10^{8.5} \text{K}$, while the Recal model and AGNdT9 uses 10^9K . A larger ΔT leads to fewer, more energetic feedback events, whereas a lower ΔT leads to more continual heating. More details in section 4.6.4 of Schaye et al. (2015).

The free parameters in the subgrid routines were tuned to match the observational constraints on the stellar mass function, reasonable galaxy sizes, and BH – stellar mass relation respectively at $z \sim 0$.

1.3 Dust in Galaxies

Dust is a key ingredient of the ISM and the galaxy physics. It plays a major role in the ISM chemistry as well as act as a catalyst for the formation the hydrogen molecule (H_2) and thus triggering star formation processes. Dust also act as a cooling channel for gas and causes metal depletion in the ISM. Dust significantly absorbs optical/UV light and re-emits it at longer wavelength. It has less impact on higher wavelengths. This re-emission causes galaxies to look redder than they actually are. Thus dust in every part of the Universe complicates the interpretation of observations carried out. It is estimated that about 30% of light in the Universe is reprocessed by dust (Bernstein et al., 2002) and it is also necessary for, and thus traces, star formation in galaxies (Sanders & Mirabel, 1996). Hence detailed knowledge of dust is required to understand galaxy evolution and study the lifecycle of the ISM. Thus it is very important to understand the physical properties that shape the observed light at different wavelengths from galaxies, to answer these questions. However, there still remains several uncertainties in our understanding of dust grain properties and their evolution (see Galliano et al., 2018; Salim & Narayanan, 2020, for a review).

Dust is thought to be comprised of micrometer sized grains which is usually compounds of carbonates, silicates and polycyclic aromatic hydrocarbons (PAHs). In order to understand how galaxies and the IGM build its dust content one should look into the various processes that create and destroy dust. Dust particles or grains arise from the evolution of stars in galaxies. They form in stellar winds as well as supernovae remnants, then mixes with the ISM. Dust grains can also form in the ISM by grain growth, through the accretion of atoms and molecules. With many high-redshift studies finding very high dust content at high-redshift ($z > 6$, e.g. Mancini et al., 2015; Knudsen et al., 2017), which is in excess of that produced solely from SN remnants and AGB stars, this has been found to be an important avenue of dust production.

The constitution of the dust grains can change across the course of time through processes such as shattering and fragmentation during grain-grain collisions, impact by high-energy photons and coagulation. They can also be destroyed by processes of thermal or kinetic sputtering, thermal evaporation as well as astration by stars. They can also be removed by outflows, if the dust is coupled to the gas. The interplay of these different phenomena on the dust grains determine how the intrinsic emission from stars, predominantly in the UV to near-IR (NIR) is

affected. This process manifests as two phenomena: *extinction* and *attenuation*. The term *extinction* refers to the amount of light lost along a line-of-sight when travelling through a dusty medium due to absorption or scattering away from the line-of-sight. The extinction of a medium at a particular wavelength is dependent on the dust grain properties such as its size and composition as well as the optical properties of the grain and it scales with the dust column density along the line-of-sight. Attenuation includes the same mechanisms as extinction, along with scattering of light back into the line-of-sight as well as contribution from unobscured stars. The effect of all this phenomena is characterised by attenuation curves.

Observational studies of the nearby Universe (our Galaxy and the Magellanic clouds) has unearthed different extinction curves (as a function of wavelength). Many of these empirical forms like the average ones obtained from Milky Way (Fitzpatrick & Massa, 2007), SMC or LMC (Weingartner & Draine, 2001), etc are used in observational and theoretical models to infer various galaxy properties. These are the basis for obtaining attenuation curves, since any change in the former would be reflected in the latter, with the effect of line-of-sight being non-trivial to include. Most of these studies have found that there is large variations even within a galaxy for various lines-of-sight, and different galaxies exhibit very different average extinction curves. Recently there have been numerous works using simulations to understand these variations, and they have been usually attributed to the varying dust contents as well as the complex star-dust geometry within galaxies (e. g. Narayanan et al., 2018; Salim & Narayanan, 2020; Liang et al., 2021).

In the last decade there has been an influx of theoretical simulations that deals directly with the dust production and destruction mechanisms within galaxies (e. g. McKinnon et al., 2016; Aoyama et al., 2017; Popping et al., 2017a; Li et al., 2019; Triani et al., 2020, in Chapter 2 we will describe the dust model implemented in L-GALAXIES SAM). Usually the large scale simulations relies on a simplified model of dust interactions usually assuming average dust properties, or no interaction with the interstellar radiation field. These works have been very useful to understand the build up and distribution of dust in galaxies through various processes that were mentioned earlier. These works use single or multiple grain sizes to track these various processes. They have shown various successes and failures when compared to observations and have also challenged our notion of nearly constant dust-to-metal ratios across galaxies as well as dust enrichment in the early Universe, with grain growth expected to have a major impact.

1.4 Observations of the High Redshift Universe

The observation of galaxies in the early Universe ($z \geq 5$) is complicated by a number of reasons: most of the galaxies are very young and have not had time to build up their stellar content and thus are faint, they can be clumpy, thus missed by observations, cosmological redshift shifts the UV-optical to the near-IR (whose measurement from the ground is complicated by atmospheric absorption). These difficulties have been overcome with the valuable help of a large number of telescopes, both ground and space based.

Ground based observatories like *Keck* (imaging, spectroscopy and integral field units in the optical to NIR), *Subaru* (imaging and spectroscopy in the optical to NIR) and *Visible and Infrared Survey Telescope for Astronomy* (*VISTA*, large area sky surveys in the NIR using *Vista InfraRed CAMera*, *VIRCAM*), etc have enabled the detection of a large number of galaxies at these high-redshifts using their UV to NIR emission as well as nebular line emission. Detections at these wavelength range have also been possible due to space based observatories like the *Hubble Space Telescope* (*HST*, observer frame $< 1.7\mu\text{m}$), *Spitzer Space Telescope* (observer frame $3 - 180\mu\text{m}$), etc. Deep imaging conducted on the now known Hubble Deep Field (HDF), Hubble Ultra Deep Field (HUDF), Hubble eXtreme Deep Field (XDF), CANDELS, etc with *HST* have revolutionised our understanding of the high-redshift Universe with the detection of 1000s of galaxies within the past 3 decades. These detections have been complemented with NIR (observed) photometry from *Spitzer*, helping us to understand the SEDs of galaxies short of the rest-frame optical.

Beyond the rest-frame UV to optical, observations in the rest-frame far-IR (FIR) have been possible in the observed frame sub-mm/mm range due to single dish telescopes and interferometers such as the *Atacama Large mm/sub-mm Array* (*ALMA*), *Giant Metrewave Radio Telescope* (*GMRT*), *James Clerk Maxwell Telescope* (*JCMT*), *Plateau de Bure Interferometer* (*PdBI*, now succeeded by the *Northern Extended Millimeter Array*, *NOEMA*), *Very Large Array* (*VLA*), etc. They have been successful in characterising the gas and dust content of galaxies at high-redshift by detecting their dust continuum emissions as well as through line emission from fine-structure transitions like [CII] ($158\mu\text{m}$) and [OIII] ($63\mu\text{m}$).

The combination of observations on these different telescopes have enabled us to partly understand the nature of these high-redshift galaxies. Even though the current dataset is small, the future is very promising with the addition of a number of telescopes such as *Euclid* (near-IR photometry and spectroscopy), *Extremely Large Telescope* (*ELT*, optical to mid-IR, providing imaging and spectroscopy), *James Webb Space Telescope* (*JWST*, near-IR to mid-IR providing imaging, grism spectroscopy and integral field unit), *Nancy Grace Roman Space Telescope* (*Roman*, photo-

metry and grism spectroscopy in the optical to near-IR), etc. With the unprecedented amount of observational data expected over the coming decade, it is crucial to have sophisticated theoretical models of galaxy formation and evolution in order to study them. Modelling these observations and providing predictions for these upcoming surveys would provide a great avenue to test the employed physics models and update our understanding.

1.4.1 Spectral Energy Distribution Modelling

The UV to IR spectral energy distribution (SED) of a galaxy encodes information about its various physical properties, with different physical processes affecting different wavelengths. These include for e. g. properties like the stellar initial mass function (IMF), stellar metallicity, star formation history (SFH), the total stellar mass, amount of gas and dust, etc. In order to obtain these physical properties, there is a need to model the emission and fit the SED, a technique commonly referred to as SED fitting. In case of theoretical simulations, one needs to go from the simulated physical properties to an SED, there is no fitting involved, but need to model the same features required in SED fitting. For modelling one needs to assume an initial mass function (IMF, in simulations this is already assumed), stellar population synthesis (SPS) model, emission and attenuation from nebular regions and photo-dissociation regions (PDRs), impact of dust attenuation, and any contribution from AGNs, etc. Now, the reliability of the derived SEDs or the physical properties is altogether a different can of worms. As alluded to in the review by [Kewley et al. \(2019\)](#), "making a model of a galaxy by studying its spectrum is like modelling an entire symphony orchestra from the noise it makes when falling downstairs".

A number of techniques of differing complexities have been adopted to build the various parts of this framework, by various groups across the globe. Here I will describe very briefly some of these building blocks. More detailed reviews can be found in [Walcher et al. \(2011\)](#); [Conroy \(2013\)](#).

1.4.1.1 Initial Mass Function

The study of the origin of the stellar mass functions and UV luminosity functions is closely tied to the Initial Mass Function (IMF). The IMF was first introduced by [Salpeter \(1955\)](#) to provide a convenient way of parameterising the number density of stars as a function of their mass for a newly formed stellar population, with the proposed form as,

$$\frac{dn}{dM} \propto M^{-\alpha}, \quad (1.9)$$

with $\alpha = -2.35$, fit for observational data of stellar masses above a few solar masses. Even after more than 60 years, this is considered pretty standard for stars above $1M_{\odot}$. This function

diverges as it approaches zero, hence there must be a break or turnover at low masses as well as the high number of sub-solar mass stars being unfeasible with recent measurements of the stellar luminosity function. The lower mass limit of the IMF is determined by what mass of the initial gas clump is enough to start collapse and start fusion, which is $\sim 0.08M_{\odot}$.

The recent forms of the IMF adopts a log-normal distribution at low masses and a power-law above a solar mass, like the Chabrier IMF (Chabrier, 2003) or represent the full mass range of the IMF as a series of power-laws, like the Kroupa IMF (Kroupa, 2001). Both these agree quite well at the high mass end, while the IMF at the low mass end is quite uncertain and subject to ongoing debate.

1.4.1.2 Stellar Population Synthesis

Stellar population Synthesis (SPS) is the process of creating a galactic spectrum through sum of the spectra of its stars, pioneered by works in the 1970s (Tinsley, 1972; Searle et al., 1973). SPS models are built on top of simple stellar populations (SSP), which describes the SED of a single coeval stellar population at a single metallicity and elemental abundances. An SSP also requires a stellar evolution theory based on isochrones or other models, stellar spectral libraries and an IMF all which can depend on the metallicity and the elemental abundances. The sum of the spectra of SSPs with various ages and metallicities integrated over their IMF, gives composite stellar populations (CSPs), and are used to create galactic SEDs. This can be represented in the following way,

$$L_{\nu}^{\text{SSP}}(t, Z) = \int_{M_{\min}}^{M_{\max}} \phi(M) L_{\nu}(M, t, Z) dM, \quad (1.10)$$

where M is the initial (zero-age main sequence) stellar mass, $\phi(M)$ is the IMF, and M_{\min} and M_{\max} are the lower and upper mass limit of the IMF. $L_{\nu}(M, t, Z)$ is the time and metallicity dependent SSP spectrum based on the SPS code. In many cosmological simulations, stars are represented by mass elements that are massive (for e. g. $\sim 10^6 M_{\odot}$ in the EAGLE reference simulation), much larger than typical star clusters. These elements are assumed to be a single SSP.

With great improvements in modelling as well as the codes being used, SPS models today can recreate broad-band UV to NIR SEDs and high-resolution spectra remarkably well. A number of SPS codes have been developed such as PEGASE (Fioc & Rocca-Volmerange, 1997), Starburst99 (Leitherer et al., 1999), BC03 (Bruzual & Charlot, 2003), Mo5 (Maraston, 2005), FSPS (Conroy et al., 2009), etc. However, it should be noted that even though significant improvements have been made in the field of stellar libraries and their evolution, there are still parts which are only weakly understood, and thus poorly treated. Some of these such as the effect of binaries (recently

BPASS models, see [Stanway & Eldridge, 2018](#), have included their effects), stellar rotation as well as rapid phases in the population such as thermally pulsing asymptotic giant branch (TP-AGB) stars, extreme horizontal branch stars (EHB), blue stragglers, etc are challenging to model and incorporate.

1.4.1.3 Nebular Emission

The ionised gas heated by young stars, produce nebular line and continuum emission which contains valuable information about the nature of these stars as well as the physical conditions in the ISM. Many prominent optical emission lines are used to estimate the density, metallicity, whether ionisation of the ISM is dominated by young stars, AGN or evolved AGB stars (see [Kewley et al., 2019](#), for a review). The near-IR instruments on *JWST* will be able to probe many of these lines up to very high-redshifts. NIRSpec instrument can be used to get deep NIR single slit, multi-object, and integral field spectroscopy from $\sim 0.7 - 5\mu\text{m}$, while the NIRISS and NIRCам instruments will provide near-IR wide field slitless spectroscopy, which is sufficient to probe many of the strong rest-frame optical emission lines at high-redshifts.

To model the emission from these regions, SPS codes (to estimate the ionising radiation) are coupled with photo-ionisation modelling codes like CLOUDY ([Ferland et al., 2017](#)) or MAPPING (Groves et al., 2004a). The effect of nebular emission is very complex, since it depends on the different assumptions used in deriving the fluxes. It is usually very important in galaxies which have low metallicity young stars, with the contribution to the broadband fluxes that can reach 20 – 60% (e. g. [Anders & Fritze-v. Alvensleben, 2003](#)). Thus it will have a larger effect at high-redshift where high SFR and low metallicity galaxies are more common. Another contributing effect is the redshifting of the rest-frame Equivalent Widths (EWs) being stretched over the broadband filters.

When modelling nebular emission, the ionising radiation from the SSP is used to create individual HII regions or birth clouds. This is usually done by characterising the region using the dimensionless ionisation parameter U_S at the Strömgren radius (R_S) with the hydrogen number density (n_H) and the number ionising photons per second (Q , obtained from the SSP),

$$U_S = \frac{Q}{4\pi R_S^2 n_H c} \text{ or} \quad (1.11)$$

$$U_S \propto (Q \epsilon^2 n_H)^{1/3}. \quad (1.12)$$

Here ϵ is the effective volume filling factor of the gas. $\epsilon^2 n_H$ encodes the approximate geometry of the region. Nebular emission has been implemented with different SPS models with varying

complexity (e. g. [Fioc & Rocca-Volmerange, 1997](#); [Gutkin et al., 2016](#); [Byler et al., 2017](#)), as well as included in post-processing of simulations (e. g. [Hirschmann et al., 2017, 2019](#); [Wilkins et al., 2020](#)).

1.4.1.4 Fine-structure transitions

In the past decade, the high-redshift Universe (even out to $z > 9$) has increasingly been probed by sub-mm/mm interferometers using emission lines in the FIR regime arising from fine-structure transitions such as [CII] (158 μ m, rest-frame), [OI] (63 μ m, rest-frame), [OIII] (88 μ m, rest-frame), etc (e. g. [Swinbank et al., 2012](#); [Capak et al., 2015](#); [Knudsen et al., 2016](#); [Hashimoto et al., 2019](#)). [CII] is the dominant cooling line in the neutral ISM ([Dalgarno & McCray, 1972](#)), and is also the best studied line among this. It is the brightest IR emission line in the spectrum of most galaxies, providing as much as 1% of the total FIR luminosity ([Stacey et al., 1991](#)). Except for sources with substantial redshifts, [CII] can be observed only from above the Earth’s atmosphere due to absorption by water vapour in the atmosphere. At high redshifts the lines are shifted towards the sub-mm bands and thus observable from ground based telescopes. Being an extremely bright line, it would be one of the first emission line to be picked up by sub-mm observations, making it an important line to detect new objects using blind surveys at high-redshift using *ALMA*.

The detection of these fine-structure lines have contributed enormously to understanding the dynamics of the gas and the star formation activity in addition to obtaining spectroscopic redshift of the galaxies in the high-redshift Universe. [CII] and [OIII] have been found to empirically correlate with the star formation rate of the galaxies in the local Universe ([De Looze et al., 2011, 2014](#)), with the relationship at high-redshift being unclear. [CII] emission can arise from nearly every phase of the ISM (due to its low ionisation potential) making its interpretation complicated, while [OIII] traces the ionised medium.

To model the emission of these lines, SPS codes are coupled with photo-ionisation modelling codes to estimate the line luminosity arising from molecular clouds or ionised regions, with the strength usually dependent upon their density, temperature, metallicity and size of the PDR. Some of these lines have already been modelled in semi-analytical models (e. g. [Popping et al., 2016](#); [Lagache et al., 2018](#); [Popping et al., 2019](#)) and hydrodynamical simulations (e. g. [Vallini et al., 2015](#); [Olsen et al., 2017](#); [Katz et al., 2017](#); [Moriwaki et al., 2018](#); [Leung et al., 2020](#)) in post-processing.

1.4.1.5 Dust Attenuation

Seminars and colloquia would have been incomplete a decade ago, without the proverbial question ‘Have you considered dust?’. Recent years have seen both observations and simulations model the effect of dust in a variety of ways with differing complexities.

Usually after modelling all the machinery described above one can get the intrinsic SED of a galaxy. In order to obtain the observed SED, a model for dust attenuation needs to be assumed. Usually the large scale hydrodynamical simulations or SAMs do not explicitly model dust production and destruction (for e.g. BLUETIDES, EAGLE, GALFORM, ILLUSTRIS-TNG, etc, unlike for e.g. SIMBA which includes a passive empirical model for dust), and use simplified recipes in post-processing to obtain the observed galaxy SED. There are a number of ways to do this.

A simple model for dust attenuation assumed in many studies is a screen or slab in front of the stellar populations. Even within this, one can introduce complexities, ranging from simple foreground screens (single attenuation across) to mixed slabs or discrete clouds (see for e.g. [Charlot & Fall, 2000](#)). This can be implemented in the following way, with observed flux at a particular wavelength ($F_{\text{obs}}(\lambda)$) expressed as

$$F_{\text{obs}}(\lambda) = F_{\text{int}}(\lambda) \exp(-\tau(\lambda)), \quad (1.13)$$

where $F_{\text{int}}(\lambda)$ and $\tau(\lambda)$ is the intrinsic flux and the optical depth at wavelength λ . $\tau(\lambda)$ can be parameterised in different ways by assuming an attenuation curve from literature, for example the Calzetti law ([Calzetti et al., 2000](#)), the Milky Way ([Fitzpatrick & Massa, 2007](#)), SMC ([Pei, 1992](#)), or time dependent attenuation model in [Charlot & Fall \(2000\)](#), etc. Screen models have also been implemented in simulations of galaxy formation, coupling the dust content to the metal content (e.g. [Clay et al., 2015](#); [Trayford et al., 2015](#)).

Instead of using a screen or slab model, it is possible to include the distribution of dust when implementing attenuation. There have been different methods that can be used to accomplish this in hydrodynamical simulations where this distribution can be inferred. A simple method used is to calculate a line-of-sight (los) attenuation model, which can be done by calculating the dust los density. In cosmological simulations which does not explicitly model dust production and destruction, dust los density can be assumed to be proportional to the metal los density. Now by combining the optical depth with an attenuation curve, the observed SED can be obtained. A drawback of such method is that it is hard to incorporate effects of scattering to and away from the los.

A computationally expensive and comprehensive way that can also capture the properties

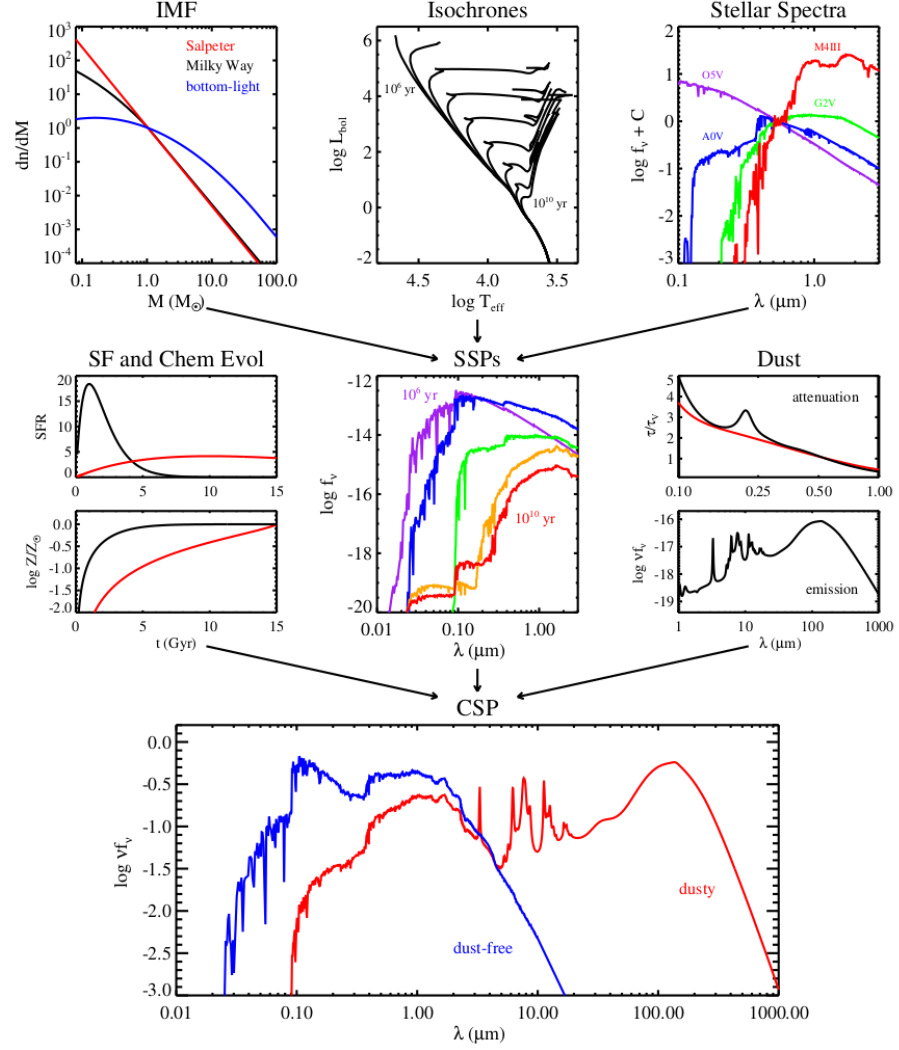


Figure 1.4: Top panel: Ingredients required to create SSPs, such as the IMF, stellar isochrones for different age and metallicities, and the associated stellar spectra. Middle panel: Ingredients required for constructing CSPs, such as the star formation histories and chemical evolution, the SSP and a model for dust attenuation. Bottom panel: Shows the final galaxy CSP before (blue) and after (red) dust attenuation. Reproduced from [Conroy \(2013\)](#).

of the dusty medium is to run radiative transfer (RT) code on the simulated galaxies. This approach can properly capture the effect of scattering, absorption and re-emission of radiation by dust grains. The last decade has seen many popular open source radiative transfer codes (e. g. SUNRISE ([Groves et al., 2004a](#)), RADMC ([Dullemond et al., 2012](#)), SKIRT ([Camps & Baes, 2015, 2020](#)), POWDERDAY ([Narayanan et al., 2021](#)), etc) being developed that can fully capture the 3D emission using Monte Carlo techniques. SKIRT is a very popular code that have been applied to post-process galaxy simulations (e. g. [Camps et al., 2016](#); [Trayford et al., 2017](#); [Ma et al., 2019](#); [Vogelsberger et al., 2020b](#)) to produce realistic SEDs. In this thesis I have used SKIRT to model the full SEDs of the most massive galaxies in the high-redshift Universe. Below I describe briefly the scheme very briefly in terms of the EAGLE simulation physics.

1.4.1.6 SKIRT Radiative Transfer code

SKIRT (Camps & Baes, 2015, 2020) as alluded to earlier is a 3D Monte Carlo dust radiative transfer code for simulating the effect of dust on radiation in astrophysical systems, mainly used by simulations of galaxy formation and evolution in post-processing. The code offers the user the ability to model the absorption and scattering of radiation by dust, computing the dust temperature across spatial resolutions, taking into account the effect of re-emission and supports CMB and stochastic heating of dust grains. The code allows you to implement various models for the dust grain mixture and different 3D geometries of the radiation sources. The array of available models can be chosen and configured into a parameter file with the help of a very user-friendly interface.

To apply SKIRT to the simulation data require inputs such as information on the galaxy stellar properties and distribution, assumptions on dust grain type and distribution (which reproduces the seen attenuation curves), number of photons required per source grid, assumptions on local thermal equilibrium, spatial grid resolution, etc as mentioned. They have been used to study the galaxy UV/IR luminosity functions, effect of dust attenuation, various dust temperatures as well as their redshift evolution and dependence on intrinsic properties. It should also be noted that there are also many caveats when using any of these methods due to the involvement of many free parameters that are usually tuned to obtain observables (e. g. luminosity functions) or by the use of some subgrid models that could be termed as black boxes (for e. g. the use of MAPPINGS III (Groves et al., 2008) for modelling young stars in SKIRT).

The details on how the EAGLE simulation data is post-processed is described extensively in Camps et al. (2016); Trayford et al. (2017), I describe very briefly some of the details used in the standard EAGLE analysis here:

1. Similar to standard EAGLE, only gas and star particles within 30 pkpc aperture centred on the galaxy centre of potential are used in the modelling.
2. Star particles are divided into old (with age > 100 Myr) and young stellar populations based on their age. The old stars are assigned GALAXEV (Bruzual & Charlot, 2003) SEDs with a Chabrier IMF. Uses information on the star particle coordinate, smoothing lengths, initial mass, metallicity and age.
3. To build star-forming regions, a resampling technique is adopted to take care of sampling issues. This is done by selecting every star particle with age less than 100 Myr as well as star-forming gas particles (gas with non-zero SFR, see Schaye et al., 2015; Crain et al., 2015,

for more details) and then resampled into a number of sub-particles with masses drawn randomly from the power-law mass distribution function,

$$\frac{dN}{dM} \propto M^{-1.8} \quad \text{with } M \in [700, 10^6] M_{\odot}. \quad (1.14)$$

These sub-particles are all assigned a formation time sampled randomly to represent their parent's SFR (in case of star particles their SFR at birth) and mass, assuming a constant SFR over a 100 Myr lifetime. Particles with formation time less than 10 Myr (typical lifetime of birth clouds) are designated as star-forming regions; those that have their formation time greater than 10 Myr are assigned as old stars; while the particles that have not yet formed are cast as gas particles. The star-forming gas particles are modelled using the MAPPINGS III SEDs which uses the Starburst99 (Leitherer et al., 1999) SPS model with the Kroupa IMF (Kroupa, 2001). This takes in information on the particle coordinate, smoothing length, metallicity, SFR, gas density at the HII region position, pressure of the ambient ISM, compactness of the HII region (Groves et al., 2008, calculated using equation 13 in) and the PDR covering fraction (f_{PDR} , a free parameter, chosen as 0.1 in the EAGLE analysis). The free parameters (see Camps et al., 2016, for more details) were chosen by comparing far-infrared mock observations of EAGLE galaxies with observations of local galaxies in the *Herschel* Reference Survey (HRS, Boselli et al., 2010).

4. 'Ghost' particles are introduced at the site of star-forming particles to account for the dust mass already associated with birth clouds when using the MAPPINGS-III SEDs. They contribute negatively to the dust mass density (see section 2.4.4 and 3.2.2 in Camps et al., 2016; Trayford et al., 2017, respectively for more details).
5. To model the dust distribution, SKIRT takes in the gas coordinate, smoothing length, gas mass, gas metallicity, SFR, temperature and a dust-to-metal ratio (DTM) as well as an upper limit on the gas temperature. EAGLE does not model dust self-consistently, hence a DTM ratio (free parameter) is assumed for the gas particles to obtain the dust mass. In the standard EAGLE analysis the dust-to-metal ratio is assumed to be 0.3 for gas particles with a temperature of less than 8000 K.
6. The number of photons per wavelength grid as well as the wavelength grid resolution is selected such that convergence is attained on using higher values.

In Chapter 5, we use a variable DTM ratio obtained using the fitting function in Vijayan et al. (2019) and changed the maximum temperature of the gas particle in which dust is not destroyed to 10^6 K.

The techniques mentioned above are all applied to simulations in post-processing. There are also codes that can couple the effects of radiation to the ISM, while a simulation is running (which is usually treated subgrid). This can make the computations very expensive. Solving the moments of RT equations (instead of Monte Carlo methods, [Dullemond et al., 2012](#), for example) have gained popularity in recent years due to their computational accuracy and efficiency (e. g. [Rosdahl & Teyssier, 2015](#); [Kannan et al., 2019](#); [Chan et al., 2021](#)). This is very remarkable, considering the fact that to properly resolve the effects of radiation on the ISM, high resolution multi-phase simulations are required. Thus these studies are usually limited to high-resolution small volume simulations (e. g. [Finlator et al., 2018](#); [Rosdahl et al., 2018](#); [Wu et al., 2019](#)). An exception to this is the CoDa I ([Ocvirk et al., 2016](#)) and CoDa II ([Ocvirk et al., 2020](#)) simulations that uses a hybrid CPU-GPU code RAMSES-CUDATON to speed up the radiative transfer calculations. The codes are also getting more sophisticated in the physics treatment, with the recent AREPO-RT ([Kannan et al., 2019](#)) code, which couples the infrared RT scheme to the semi-empirical dust model of [McKinnon et al. \(2016\)](#); [McKinnon et al. \(2018\)](#) implemented in AREPO.

Figure 1.4 (reproduced from [Conroy, 2013](#)) gives an overview on how the different frameworks for building the galaxy SED are connected with each other.

1.5 Thesis Overview

I have discussed briefly some of the framework that will be utilised in this thesis. With the unprecedented volume of observational data expected in the coming decade, theoretical models need to keep up to better understand the astrophysical processes in the early Universe. The following chapters will explore the evolution of dust in galaxies and its effect on the observational spaces using semi-analytical models and hydrodynamical simulations. All the observables from the simulations are compared to the available observations in that space. In the 2nd Chapter, we extend the L-GALAXIES SAM to incorporate a model for dust production and evolution. In the 3rd Chapter, the suite of hydrodynamical simulations termed First Light And Reionisation Epoch Simulations (FLARES), re-simulating a range of overdensities in the EoR using the EAGLE simulation physics is introduced. We discuss the motivation to perform simulations with the employed technique. In the 4th Chapter, we implement a line-of-sight dust attenuation model in FLARES to study the photometric properties of the galaxies in the EoR and comparing it to current observational constraints. In the 5th Chapter using the radiative transfer code SKIRT, we post-process the most massive galaxies in FLARES to derive their full SEDs and explore their dust driven properties. In Chapter 6, I present my conclusions and plans for future work.

DETAILED DUST MODELLING IN THE L-GALAXIES

2

SEMI-ANALYTIC MODEL OF GALAXY FORMATION

2.1 Introduction

Dust has a major impact on the observed properties of galaxies with almost 30% of all photons in the Universe having been reprocessed by dust grains at some point in their lifetime (Bernstein et al., 2002). These grains can form in the stellar winds around AGB and other evolved stars, in supernovae remnants (SNR), and can grow *in situ* within molecular clouds. Processes that destroy or alter dust grains include shock heating by supernovae, photo-evaporation and chemical explosions (De Boer et al., 1987; Savage & Sembach, 1996). The dust content of a galaxy thus depends in a complex way upon the evolutionary history of its interstellar medium.

The purpose of this work is to implement a model for dust growth and destruction within the L-GALAXIES semi-analytic model in order to investigate the evolution of the dust content of galaxies, with particular regard to the high-redshift Universe.

2.1.1 Dust production and destruction

The stellar sources of dust are, in order of importance, type II SNR, AGB stars and type Ia SNR. These dust yields are dependent upon the age and metallicity of the stellar populations. For SNR we use the prescription of Zhukovska et al. (2008) and for AGB stars the tables of Ferrarotti & Gail (2006) – this is described in detail in Section 2.3.1 below. We note that at very high redshifts, $z \gtrsim 6$ observations in the far-infrared have started to identify dust masses substantially in excess

of the amount formed from SNR and AGB stars (e.g. [Mancini et al., 2015](#); [da Cunha et al., 2015](#)). It is possible, therefore, that the dust yields may be higher at earlier times, perhaps due to higher survival rates of dust produced in SNR (e.g. [Dwek et al., 2014](#)) – we do not consider that here.

It is now generally accepted that, at later times, the dominant source of dust in the Universe is grain growth inside molecular clouds (e.g. [Mattsson, 2015](#)). Our dust growth model, described in Section 2.3.2, builds on that of [Zhukovska et al. \(2008, hereafter ZGTo8\)](#) and [Popping et al. \(2017a, hereafter PSG17\)](#). Unlike earlier works, we use a variable limit for the fraction of an element that can be locked up in dust, motivated by the chemistry of the ISM, and we explicitly follow the dust growth in molecular clouds and the diffuse inter-cloud medium separately, finding that the two can be quite different in certain regimes.

Dust is destroyed by sputtering at high temperatures. In our model, we follow the prescription of [McKee \(1989\)](#) for dust destruction in SNR, described in Section 2.3.3, and we consider dust to be instantly destroyed if it is reheated out of the cold ISM to join the hot corona of the galaxy. We ignore other processes, such as interaction with cosmic rays, or ejection from the cold ISM by feedback from an active galactic nucleus – we will show in Section 2.5.3 that we have an excess of dust in massive galaxies at low redshift and this may be one possible cause of that.

2.1.2 Previous modelling

In recent years, detailed chemical enrichment models have been implemented into both semi-analytic models (SAMs, e.g. [Arrigoni et al., 2010](#); [Yates et al., 2013](#); [De Lucia et al., 2014](#)), and hydrodynamical simulations (e.g. [Wiersma et al., 2009b](#); [Vogelsberger et al., 2013](#); [Pillepich et al., 2018](#)), and a detailed modelling of the dust chemistry is the natural next step. Lately there have been works that incorporated dust models in simulations.

The current efforts of modelling dust in semi-analytic models and hydrodynamic simulations builds heavily upon the initial ‘one-zone’ models, first implemented in [Dwek \(1998\)](#) and followed up by [Inoue \(2003\)](#), [Morgan & Edmunds \(2003\)](#), [ZGTo8](#). The most detailed semi-analytic (SA) work, which we use as a basis for our own modelling, is that of [PSG17](#), which uses the SANTACRUZ ([Somerville & Primack, 1999](#)) SA model. Their model was run on a grid of haloes for a range of virial masses with trees created using the extended Press Schechter formalism; whereas our model uses the full set of trees from the relatively low-resolution but cosmologically representative Millennium ([Springel et al., 2005a](#)), and the higher-resolution Millennium II ([Boylan-Kolchin et al., 2009](#)) simulations (hereafter MR and MR II respectively). Where appropriate, we will make comparison to [PSG17](#) in the results presented below.

Recent studies (e. g. Bekki, 2013; Mancini et al., 2016; McKinnon et al., 2016, 2017; McKinnon et al., 2018; Aoyama et al., 2017; Hirashita et al., 2018; Gjergo et al., 2018; Davé et al., 2019; Hou et al., 2019) have implemented mechanisms for tracking dust production and destruction in hydrodynamical simulations. McKinnon et al. (2016, 2017) implemented a simplified dust model in the moving mesh code AREPO to investigate dust formation in a diverse sample of galaxies, accounting for thermal sputtering of grains. Their model gives results in rough agreement at low redshifts for the dust mass function, cosmic dust density and the mean surface density profiles. In McKinnon et al. (2018), the model was improved to track the dynamical motion and grain-size evolution of interstellar dust grains. They predict attenuation curves for galaxies which show large offsets from the observed ones. Aoyama et al. (2017); Hirashita et al. (2018); Hou et al. (2019) considered a simplified model of dust grain size distribution by representing the entire range of grain sizes with large and small grains. They find the assumption of a fixed dust-to-gas (DTG) ratio to break down for galaxies older than 0.2 Gigayears (Gyrs) with grain growth through accretion contributing to a non-linear rise.

2.1.3 Observational summary

To compare simulations with observational data, it is important to understand how observers calculate the dust properties of their galaxy populations. Derivations of physical dust quantities are generally done using spectral energy distribution (SED) modelling. Many observational studies of dust mass (e.g. Casey et al., 2014; Clemens et al., 2013; Vlahakis et al., 2005) in galaxies fit single or multiple greybodies to galaxy SEDs by assuming an emissivity index, β and a dust temperature, T_d , which is quite useful when the available data is limited. More complicated models can also take into account microscopic dust properties, such as the composition and grain size (Zubko et al., 2004). These models also typically assume that the properties and conditions are uniform throughout the galaxy (Rémy-Ruyer et al., 2015), or that the properties in all galaxies at all times are the same as in the local Universe (Santini et al., 2014). For all these reasons, it should be appreciated that measurements of dust mass come with large systematic uncertainties up to a factor of 2-3 (Galliano, F. et al. 2011; Dale et al. 2012).

At higher redshifts, far infrared (FIR), millimetre (mm) and sub-millimetre (sub-mm) observations are generally only possible in extreme galaxies, such as those undergoing starbursts or heavy AGN activity. Sub-mm and mm observations have been shown to be powerful tools in determining how dust and gas are evolving in high-redshift galaxies, with molecular transitions such as CO and the continuum emission used to determine the properties of gas (e.g. Greve et al., 2005; Tacconi et al., 2006; Scott et al., 2011) and dust (e.g. da Cunha et al., 2008) respectively. Sub-

mm observations are extremely good at tracing the cold dust component of the galaxy which usually dominates the dust mass. *ALMA* observations have been instrumental to systematically map the dust continuum (e. g. [Hodge et al., 2013](#); [Scoville et al., 2016](#); [Dunlop et al., 2017](#); [Franco et al., 2018](#)) and in some cases, where multi-wavelength data is available, the dust content of galaxies at redshifts of $2 - 4$ (e. g. [da Cunha et al., 2015](#)). Further complications arise from the further heating of dust at higher redshifts due to the CMB ([da Cunha et al., 2013](#)), and the lack of many observational data points in the FIR means that a dust temperature can not be calculated from the SED and one must be assumed. The assumption of a dust temperature can lead to differing dust masses by up to an order of magnitude ([Schaerer et al., 2015](#)).

The observational study of local galaxies by [Rémy-Ruyer et al. \(2014\)](#) found that the dust-to-metal ratio (DTM) is approximately constant in the majority of galaxies. However, at low metallicities, the ratio decreases, suggesting that dust destruction wins out over grain growth. Also at low redshift, [De Vis, P. et al. \(2019\)](#) found that the DTM ratio of DustPedia galaxies (see [Davies et al., 2017](#)) increases as galaxies age, before becoming approximately constant once the gas fraction drops below 60 %. For galaxies at higher redshifts ($z > 1$), the DTM ratio is seen to increase with metallicity over the broad redshift range of $2 \lesssim z \lesssim 5$ (e. g. [De Cia et al., 2016](#); [Wiseman et al., 2017](#)), again suggesting that a significant amount of dust is formed due to *in situ* grain growth in the ISM.

There also have been detections from deep *ALMA* and *PdBI* observations of galaxies at extremely high redshifts ($z > 6$) with large reservoirs of dust ($> 10^8 M_\odot$, e. g. [Mortlock et al., 2011](#); [Venemans et al., 2012](#); [Watson et al., 2015](#); [da Cunha et al., 2015](#)). Models to reproduce these (e. g. [Michałowski, 2015](#); [Mancini et al., 2015](#)) require either enhanced dust production from supernovae and AGB stars (and reduced destruction by the former), or very rapid dust production soon after chemical enrichment, suggesting very short grain growth timescales in these metal-poor environments. We will look at all these aspects of the dust evolution paradigm in the following sections.

2.1.4 Structure of the paper

This paper is structured as follows: in Section 2.2 we describe briefly the L-GALAXIES SAM and some of the key ingredients that have been incorporated, including the new two phase-model of cold ISM. In Section 2.3 we introduce our dust model and describe how it is implemented. We present our results on dust growth in Section 2.4, and of the dust content of galaxies in Section 2.5. Finally, we present our conclusions in Section 2.6.

Throughout this paper we adopt the initial mass function (IMF) of [Chabrier \(2003\)](#), assume the cosmological parameters derived by *Planck* ([Planck Collaboration et al., 2014](#)) and use a solar metallicity value, $Z_{\odot} = 0.0134$ ([Asplund et al., 2009](#)).

2.2 The Model

L-GALAXIES ([Henriques et al., 2015](#), and references therein, hereafter [HWT15](#)), has been developed over the years to include most of the relevant processes that affect galaxy evolution, also refer to §1.2.2.1 for more details. In this work we use a modified version of that model which includes: detailed chemical enrichment (Section 2.2.1); the differentiation of molecular and diffuse atomic phases in the cold gas (see Section 2.2.2); and the detailed dust model introduced in this paper (see Section 2.3). We highlight the changes relevant to our dust model below. An overview of all the physics contained within the [HWT15](#) version of the model can be found in the appendix of that paper.

The main non-standard symbols used in our model are:

- μ – fraction of the cold ISM that is in molecular clouds;
- f – fraction of metals within molecular clouds which condenses into dust;
- g – fraction of metals within the diffuse inter-cloud medium which condenses into dust.

When describing the dust content, we use the following subscripts:

- d – total amount of dust;
- j – elements;
- x – dust species.

2.2.1 Detailed chemical enrichment

Many galaxy formation models use an instantaneous recycling approximation that assumes that stars pollute their environments with metals the moment they are born. Given the long lifetimes of low-mass stars, this will introduce too many metals (and thus too much dust) at very early times. The detailed chemical enrichment model used here ([Yates et al., 2013](#)) only injects metals into the environment at the end of a star’s life. The model takes the metal production rate from stellar mass and metallicity dependent yield tables for type-II supernovae ([Portinari et al., 1998](#)), type-Ia supernovae ([Thielemann et al., 2003](#)), and AGB stellar winds ([Marigo, 2001](#)).

As discussed in [Yates et al. \(2013\)](#), we follow the prescription of [Tinsley \(1980\)](#) for the total

rate of metals j ejected by a stellar population at a time t :

$$\dot{M}_j(t) = \int_{M(t)}^{M_{\text{up}}} M_j(M, Z_0) \psi(t - \tau_M) \phi(M) dM. \quad (2.1)$$

Here $M_j(M, Z_0)$ is the mass of metals released by a star of mass M and initial metallicity Z_0 , $\psi(t - \tau_M)$ is the star formation rate at the time of the star's birth, and $\phi(M)$ represents the normalised initial mass function (IMF) by number. The lower limit of the integration, $M(t)$, is the mass of a star with a lifetime t (which would be the lowest mass possible to have died by this time), and the upper limit, M_{up} , is the highest mass star considered in this work, which is $120M_{\odot}$.

The stellar lifetimes used in the chemical enrichment calculations are taken from the [Portinari et al. \(1998\)](#) mass and metallicity-dependent tables. These provide the lifetime of stars of mass $0.6 \leq M/M_{\odot} \leq 120$ and for five different metallicities ranging from $Z = 0.0004$ to 0.05 .

With this chemical enrichment model incorporated, L-GALAXIES is able to simultaneously reproduce a range of observational data at low redshift, including the mass-metallicity relation for star-forming galaxies, the abundance distributions in the Milky Way stellar disc, the alpha enhancements in the stellar populations of early-type galaxies, and the iron content of the hot intra-cluster medium (see [Yates et al. 2013, 2017](#)).

2.2.2 Molecular gas

The standard L-GALAXIES model does not differentiate between atomic and molecular hydrogen in the cold ISM. To model this, we implement the molecular hydrogen prescription used in [Fu et al. \(2013\)](#) to split the cold gas medium into two components - the diffuse ISM and molecular clouds, based on the fitting equations in [McKee & Krumholz \(2009\)](#). In that model, the molecular gas fraction μ is given by

$$\mu = \begin{cases} \frac{4 - 2s}{4 + s}, & s < 2; \\ 0, & s \geq 2. \end{cases} \quad (2.2)$$

The parameter s in Equation 2.2 is defined as

$$s = \frac{\ln(1 + 0.6\chi + 0.01\chi^2)}{0.6\tau_c} \quad (2.3)$$

where $\chi = 0.76(1 + 3.1Z'^{0.365})$ and $\tau_c = 0.066 (\Sigma_{\text{comp}}/M_{\odot}\text{pc}^{-2}) Z'$, with $Z' = Z_{\text{gas}}/Z_{\odot}$ being the gas-phase metallicity (including metals locked up in dust) relative to the solar value. Also,

$$\Sigma_{\text{comp}} = c_{\text{f}} \Sigma_{\text{gas}} \quad (2.4)$$

where $\Sigma_{\text{gas}} = M_{\text{cold}}/\pi r_{\text{d}}^2$ is the surface density, r_{d} is the galaxy disk scale length, and c_{f} is a metallicity-dependent clumping factor given by

$$c_{\text{f}} = \begin{cases} 0.01^{-0.7}, & Z' \leq 0.01; \\ Z'^{-0.7}, & 0.01 < Z' < 1; \\ 1, & Z' \geq 1. \end{cases} \quad (2.5)$$

which is meant to account for starburst systems in low-metallicity dwarf galaxies.

In our new model, supernovae and stellar winds are assumed to inject a fraction $(1 - \mu)$ of their metal and dust into the diffuse component and a fraction μ into the molecular cloud component. However, star formation and dust growth on grains occurs only in molecular clouds.

We also note that our results remain unchanged on using the molecular hydrogen partitioning recipe used in [Martindale et al. \(2017\)](#) implementing a partitioning based on the mid-plane hydrostatic pressure in the galactic disc from [Blitz & Rosolowsky \(2006\)](#).

2.3 Detailed dust model

In this section, we describe the new detailed dust model that we have incorporated into L-GALAXIES. Our model traces the three dominant sources of dust production in the Universe; injection by type Ia and type II supernovae, stellar winds from AGB stars, and the growth of dust within molecular clouds. We also implement a model of dust destruction induced by supernovae shocks and gas heating. We make the assumption that dust grains only reside within the cold ISM, as the temperature in the hot circumgalactic and intra-cluster media around galaxies is sufficiently high that dust grains will be rapidly destroyed in those gas phases. This is an oversimplification as dust is observed in both the CGM (e. g. [Peek et al., 2015](#)) and ICM (e. g. [Gutiérrez & López-Corredoira, 2014](#)). [Tsai & Mathews \(1995\)](#) adopted an analytic form for the decrease in the dust grain radius in the hot phase. The sputtering timescale derived from this (used in other studies, e. g. [McKinnon et al. 2017](#); [Hirashita et al. 2018](#)) can vary between 1 Myr - 10 Gyr depending on the temperature and the density of the hot phase. Since the sputtering timescales of dust in the hot phase is strongly dependent on the assumed model, we do not consider that here and focus on the dust content of the ISM. This aspect will be revisited in a future work.

The dust production rate of a galaxy is therefore

$$\dot{M}_{\text{d}}(t) = \dot{M}_{\text{d},\text{inj}} + \dot{M}_{\text{d},\text{grown}} - \dot{M}_{\text{d},\text{dest}} - \dot{M}_{\text{d},\text{trans}}, \quad (2.6)$$

where $\dot{M}_{\text{d,inj}}$ is the dust yield rate from stellar sources (supernovae and AGB stars), $\dot{M}_{\text{d,grown}}$ is rate of dust growth in molecular clouds, $\dot{M}_{\text{d,dest}}$ is the dust destruction rate, and $\dot{M}_{\text{d,trans}}$ is the rate at which dust is transferred out of the cold ISM through processes such as star formation or mergers. We discuss each of these processes in more detail below.

2.3.1 Supernova and AGB dust yields

By analogy to Equation 2.1 we have

$$\dot{M}_{\text{d,inj}} = \int_{M(t)}^{M_{\text{up}}} M_{\text{d}}(M, Z_0) \psi(t - \tau_M) \phi(M) dM, \quad (2.7)$$

where $M_{\text{d}}(M, Z_0)$ is the mass of dust produced by a star of mass M and initial metallicity Z_0 , and the other parameters are as described in Section 2.2.1. We apply this equation for both AGB winds from lower-mass stars and for supernovae.

The mass of dust produced by a low mass star of given mass and metallicity (i.e. AGB stars) is taken from the tables of [Ferrarotti & Gail \(2006\)](#). In this case, the upper limit of the integral is the maximum possible mass for an AGB star, which is about $8M_{\odot}$.

For supernovae, we follow the prescription laid out in [ZGT08](#). There it is assumed that the mass of dust formed in a supernova remnant is proportional to the total mass return of the key element required to form that particular type of dust. The four types of dust they consider are silicates, carbon, iron, and silicon carbides, where the key element that comprises each species is Si or Mg, C, Fe, and Si, respectively.

We use the following equation to govern the production rate of dust formed by supernovae for the four separate dust species (denoted by a subscript x) that we consider:

$$\dot{M}_x = \eta_x \dot{M}_j \frac{A_x}{A_j}, \quad (2.8)$$

where \dot{M}_j is the mass return rate of the key element, which we obtain from our detailed chemical enrichment model as described in Section 2.2.1, and A_x and A_j are the atomic weights of the dust species and key element, respectively. The condensation efficiency parameter, η_x , is used for converting a specific element into dust, as estimated from observations of local supernovae remnants. These efficiency parameters are defined considering the effects of the reverse shock and are therefore smaller than they would be for initial dust condensation.

We apply Equation 2.8 to all four different dust species for type II supernovae, and for iron-based dust from type Ia supernovae. The values of the parameters that we use are given in Table 2.1.

Table 2.1: The conversion efficiencies used for the production of dust grains in supernovae remnants based on the mass return of key metals. The efficiencies have been adopted from ZGT08.

	Dust Species (x)			
	silicates	carbon	iron	SiC
$\eta_{x,\text{SNII}}$	0.00035	0.15	0.001	0.0003
$\eta_{x,\text{SNIA}}$	0.0	0.0	0.005	0.0
A_x	121.4	12.01	55.85	40.10
	Key Element (j)			
	Si / Mg	C	Fe	Si
A_j	28.08 / 24.31	12.01	55.85	28.09

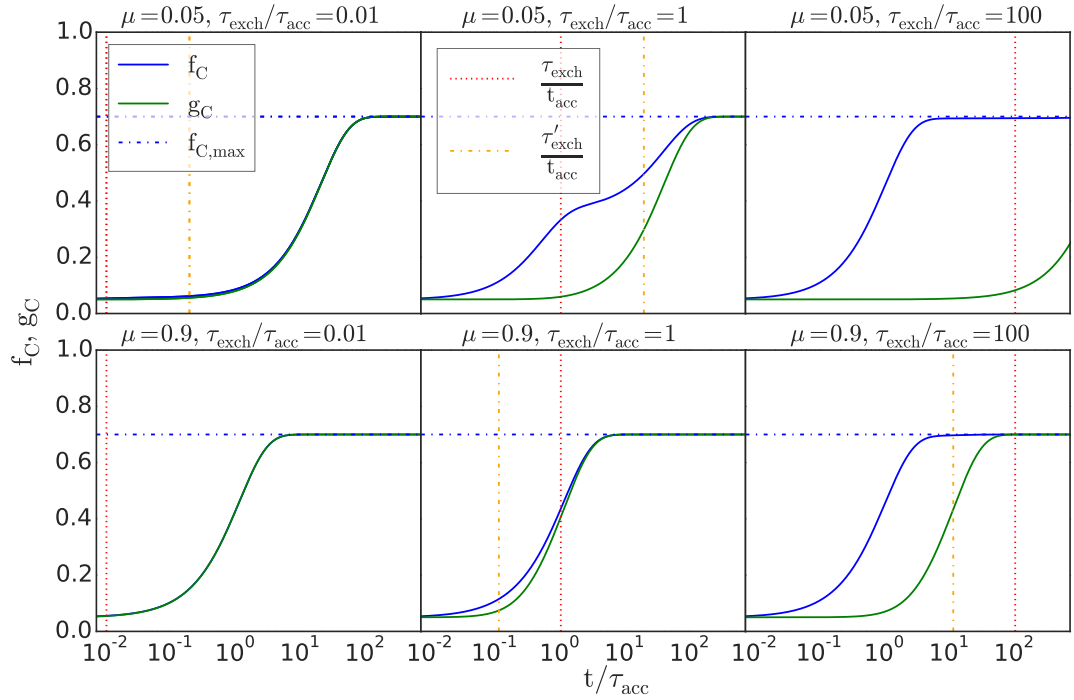


Figure 2.1: Evolution of the fractions f_C and g_C for different values of τ_{exch} and μ with an initial dust fraction $f_C = g_C = 0.05$. These plots are valid for a constant value of τ_{acc} , which in our model decreases with the production of more dust, speeding up the saturation of the two fractions. The horizontal dot-dashed line represents the maximum permissible condensation value, fixed here at 0.7 for carbon. The vertical lines show the ratios of the accretion and exchange timescales.

2.3.2 Grain growth in molecular clouds

A complete model for grain growth would consider how the accretion of different elements varies with different grain sizes, shapes, compositions and grain chemistry, but this would become very complicated. Here, we follow PSG17 in adopting a simpler model in which grain growth inside molecular clouds occurs on a timescale referred to as the accretion timescale (τ_{acc}), and exchange of materials between the molecular clouds and the diffuse media is governed by an

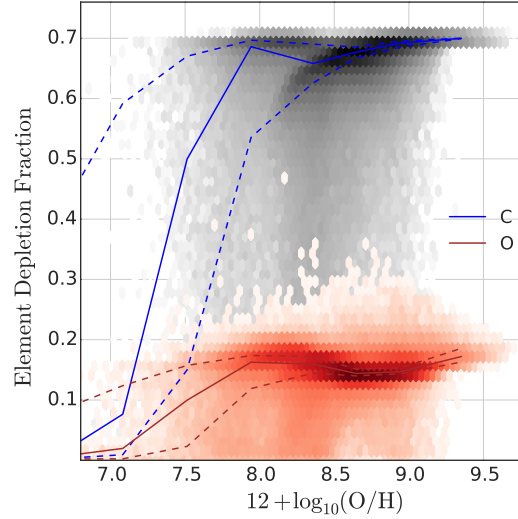


Figure 2.2: Carbon and oxygen depletion fractions plotted against the total cold-gas metallicity for $z = 0$. Blue and brown lines denote the median result from galaxies in our model, with the dashed lines denoting the 84 and 16 percentiles.

exchange timescale (τ_{exch}) which is also the average lifetime of molecular clouds and is set to 10 Myr (Zhukovska, 2014).

For each element j in the molecular-cloud component of the ISM, we set a maximum condensation fraction that can be locked up in dust, $f_{j,\text{max}}$. There is also an implicit maximum g_j , which is set by the $f_{j,\text{max}}$ in the molecular clouds. This we fix at unity for the refractory elements Mg, Si, Ca, and Fe, while for N, Ne and S it is set to 0. Neon is unreactive, nitrogen is mostly bound up in volatile gases and sulphur shows little or no incorporation into dust grains (Jones, 2000). In the case of carbon and oxygen, we follow ZGT08 to estimate $f_{j,\text{max}}$. Some carbon is locked up as CO in the molecular clouds and thus not available for grain growth. Observations estimate the fraction of carbon that is locked up as CO inside molecular clouds to be around 20-40 % (Irvine et al., 1987; van Dishoeck & Blake, 1998). In our model we fix this at 30 %, thus setting $f_{\text{C};\text{max}} = 0.7$. In the case of oxygen, we assume it is present in dust in the form of metal oxides. Thus, the maximum fraction of available oxygen is set by the amount of other elements present to form these compounds, which are silicates and iron oxides in our model. Following ZGT08, we adopt olivine ($[\text{Mg}_y\text{Fe}_{1-y}]_2\text{SiO}_4$) and pyroxene ($\text{Mg}_y\text{Fe}_{1-y}\text{SiO}_3$) as the major silicate compounds in the ISM in the ratio 32:68; here we take $y = 0.8$. In the case of iron oxides, we assume hematite (Fe_2O_3) and magnetite (Fe_3O_4) are the major compounds, contributing equally towards dust growth. Thus for oxygen the maximum condensed fraction in molecular clouds depends on the chemical composition.

Grain growth is then implemented by solving the following pair of coupled differential equa-

tions at each timestep within the simulation:

$$\frac{df_j}{dt} = \frac{f_{j,\max} - f_j}{\tau_{\text{acc}}} + \frac{g_j - f_j}{\tau_{\text{exch}}}, \quad (2.9)$$

$$\frac{dg_j}{dt} = \frac{f_j - g_j}{\tau_{\text{exch}}} \frac{\mu}{1 - \mu} = \frac{f_j - g_j}{\tau'_{\text{exch}}} \quad (2.10)$$

where f_j and g_j are the condensation fractions of element j in dust within the molecular clouds and the diffuse medium, respectively, and τ'_{exch} is the effective exchange timescale over which all the ISM in a galaxy is cycled through molecular clouds (see [Zhukovska 2014](#)).

Figure 2.1 shows how the condensation fractions f_j and g_j evolve for the particular case of carbon. Columns show different $\tau_{\text{exch}}/\tau_{\text{acc}}$ ratios, and rows show different molecular gas fractions, μ . For values of $\tau_{\text{exch}} \ll \tau_{\text{acc}}$, the condensation fractions evolve similarly for both high and low μ . For $\tau_{\text{exch}} \gg \tau_{\text{acc}}$, f_C saturates at $f_{C,\max}$ relatively quickly. However, g_C takes a much longer time to reach its maximum allowed value, with its evolution being particularly slow in regions with low μ (i. e. dominated by diffuse gas).

Because dust catalyses the formation of other dust, we use the following expression for the accretion timescale, which differs from some of the expressions used in previous studies in that it uses the dust mass instead of the metal mass in the denominator:

$$\tau_{\text{acc}} = \tau_{\text{acc},0} \times \left(\frac{\text{Total mass in clouds}}{\text{Mass of dust in clouds}} \right) \quad (2.11)$$

We require a short cooling time, $\tau_{\text{acc},0} \lesssim 5 \times 10^4$ yr to match the high dust masses observed at high redshift, and we adopt this as our canonical value (note that this is lower than the 15 Myr used in [PSG17](#) because of our use of dust fraction rather than metallicity in the growth equation). The impact of varying the value of $\tau_{\text{acc},0}$, as well as the evolution of τ_{acc} with redshift, is discussed in Appendix 2.A.

We also show in Figure 2.2 the depletion fraction i. e. $M_{j,\text{dust}}/(M_{j,\text{cold}} + M_{j,\text{dust}})$ against total gas-phase ISM metallicity in our model for the case of carbon and oxygen. We find that the typical carbon depletion fraction increases over cosmic time, whereas the typical oxygen depletion fraction maintains a value of 0.1 - 0.2 below $z \sim 4$. Our values are comparable to those adopted by emission-line modelling studies (e. g. [Groves et al., 2004b](#); [Gutkin et al., 2016](#)), and our model reproduces the expectation that oxygen has a relatively low depletion onto dust grains (e. g. [Jones, 2000](#); [Jenkins, 2009](#)).

2.3.3 Dust destruction

We implement a model of dust destruction due to the effects of supernovae induced shock waves following the prescription of [McKee \(1989\)](#):

$$\dot{M}_{d,\text{dest}} = \frac{M_d}{\tau_{\text{dest}}}, \quad (2.12)$$

where τ_{dest} is the timescale for destruction of dust.

$$\tau_{\text{dest}} = \frac{M_{\text{cold}}}{M_{\text{cleared}} f_{\text{SN}} R_{\text{SN}}} \quad (2.13)$$

where M_{cold} is the mass of the cold ISM in the galaxy, and R_{SN} the rate of supernovae type II and type Ia going off in the stellar disk, which we directly model. The other two quantities are parameters of the model: M_{cleared} is the amount of cold gas that is totally cleared of dust by an average supernovae which we fix at a lower estimate from [Hu et al. \(2019\)](#) of $1200 M_{\odot}$; f_{SN} accounts for the effects of correlated supernovae and supernovae out of the plane of the galaxy, and is set to 0.36 ([McKee, 1989](#); [Zhukovska & Henning, 2013](#); [Lakićević et al., 2015](#)). For a galaxy of similar stellar and cold-gas mass to the Milky Way, this formalism returns a τ_{dest} in good agreement with the estimates obtained by [Hu et al. \(2019\)](#) for their hydrodynamical simulations of the multiphase ISM in the solar neighbourhood.

We assume that the destruction mechanisms act equally on all types and locations of dust, so that Equation 2.12 can be applied equally to all dust species. We do not consider dust destruction due to UV radiation, cosmic rays or grain-grain collisions.

2.3.4 Dust transfer

In this section, we briefly describe the other physical processes within L-GALAXIES that act on material within the cold gas phase and thus impact the dust content of galaxies.

2.3.4.1 Star formation

Stars form from the material present in their birth clouds. We therefore transfer the dust within molecular clouds to the stellar component in proportion to the total mass of stars formed:

$$\dot{M}_d = -\frac{M_d}{M_{\text{cloud}}} \dot{M}_* \quad (2.14)$$

where M_d is the mass of dust within, and M_{cloud} the total mass of, the molecular clouds, and \dot{M}_* is the star formation rate. It should be noted that the star formation prescription is the same as in [Henriques et al. \(2015\)](#).

2.3.4.2 Mergers

L-GALAXIES has separate prescriptions for minor and major mergers. In a major merger, the gas discs of the two progenitor galaxies are assumed to be completely removed through merger-induced star formation and the associated galactic winds driven by supernovae. As we made the assumption that dust can only exist within the ISM, this effectively destroys the dust.

In a minor merger, the disc of the larger galaxy survives and the cold gas component of the smaller galaxy is accreted onto it. In this case, we assume the dust components of the two merging galaxies survive and are placed into the respective disc component of the more massive galaxy.

2.3.4.3 Other dust destruction mechanisms

There are several other mechanisms, such as reheating or cooling, that transfer dust between different gas phases in a galaxy, such as when supernovae heat up cold gas. Whenever any dust is transferred out of the ISM within our model, we destroy that dust and return it to its metal components. Since we assume dust is completely destroyed in the hot phase, no dust gets transferred from hot to cold phase – this will not significantly alter our results, as there is already a strong equilibrium between the rate of dust production and destruction in the ISM in our current formalism. We direct the reader to the appendix of [HWT15](#) for a complete description of all the processes that affect the gas phases.

2.4 Results: Dust Growth

In this section, we begin to present some of the results of our model regarding the nature and efficiency of dust growth; in the next section, we will look at the resultant dust content of galaxies. We run the model using the dark matter subhalo trees from the Millennium (hereafter MR, [Springel et al. 2005a](#)) and Millennium-II (hereafter MR II, [Boylan-Kolchin et al. 2009](#)) N-body simulations of hierarchical structure formation, in order to test our model on a cosmological volume of galaxies, applying a stellar mass selection cut in the respective simulations. Galaxies below/-above a stellar mass of $10^9 M_\odot$ are selected from from MR II/MR, respectively.¹ The disjoint median lines and hex density on the plots that follow can be attributed to the different volumes of the two simulations. The analysis is restricted to central galaxies (the most massive galaxy inside the halo virial radius), unless stated otherwise.

¹ The precise choice is unimportant as the two agree over approximately a decade in the stellar mass function.

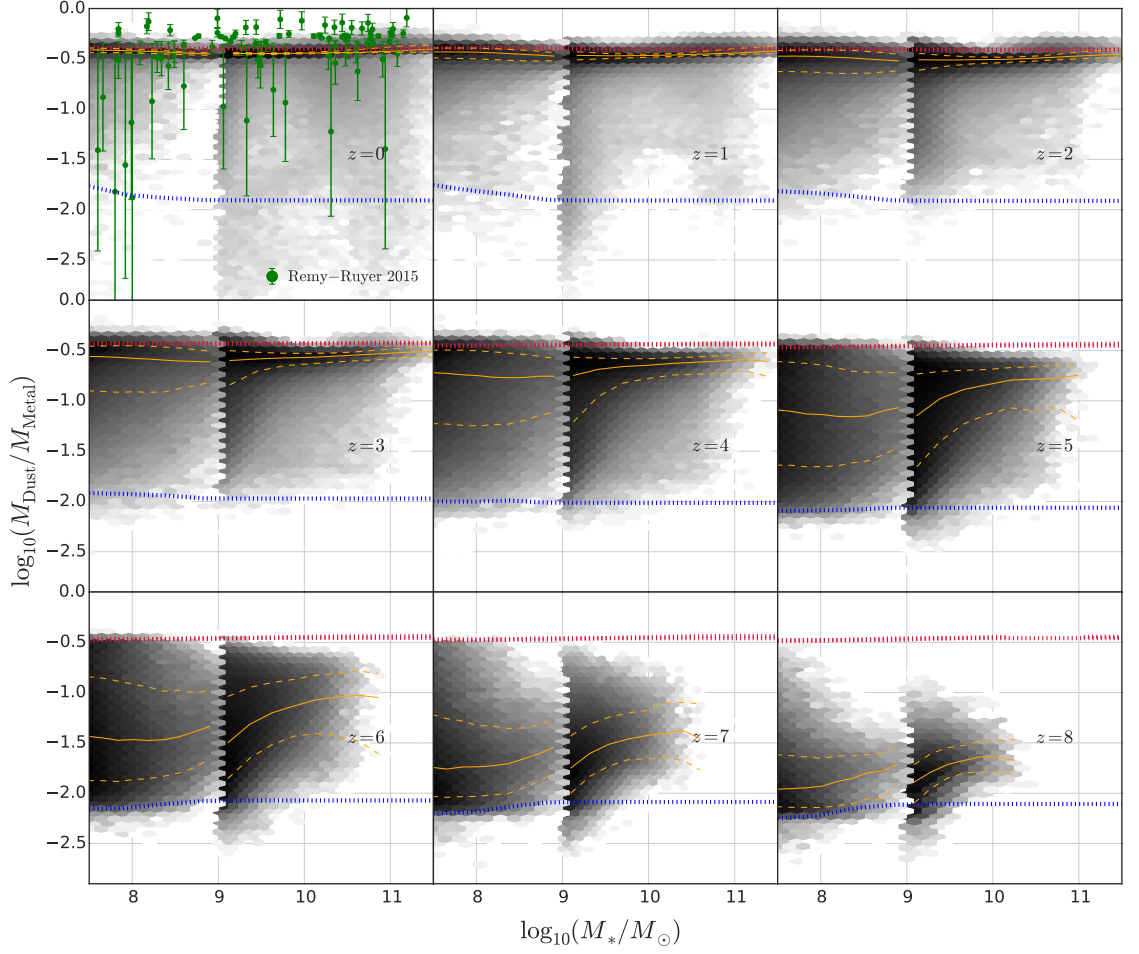


Figure 2.3: The dust-to-metal ratio as a function of stellar mass from $z = 0 - 8$. The orange line shows the median result from galaxies in our model, with the dashed lines denoting the 84 and 16 percentiles. The red dotted line represents the saturation limit calculated from average metal abundances in the model while the blue dotted line is the median DTM ratio obtained from stellar injection alone. Green points show the observational constraints from [Rémy-Ruyer et al. \(2015\)](#).

2.4.1 Dust-to-Metal (DTM) ratio

The most fundamental diagnostic and test of our model is the dust-to-metal (DTM, $M_{\text{dust}}/M_{\text{metals}} + M_{\text{dust}}$) ratio which measures the efficiency with which metals are converted in to dust.

2.4.1.1 DTM versus stellar mass

Figure 2.3 shows how the DTM ratio varies with stellar mass, with the green coloured observational data points taken from [Rémy-Ruyer et al. \(2015\)](#). The solid line shows the median result of the galaxies in our model, while the dashed lines show the 84 and 16 percentiles. The hex diagram in grey shows a 2D density distribution of galaxies in our model. The dotted, red line

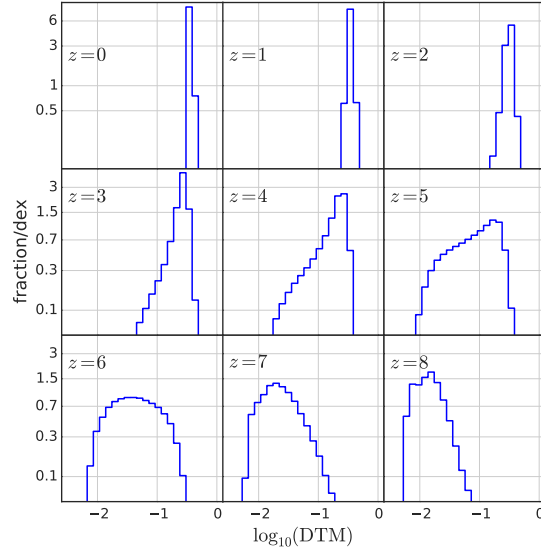


Figure 2.4: The distribution of DTM ratios for various redshifts from 0 – 8. The peak of the distribution clearly shifts from low to high values over cosmic time.

in the plot shows the maximum possible DTM ratio in our model (for the median metallicity), assuming that grain growth has saturated (i.e. $f_j = f_{j,\text{max}}$ for every element). The blue dotted line shows the median DTM ratio obtained from stellar dust production mechanisms alone. The slight displacement of the median DTM ratio below the saturation value at low redshift is due to dust destruction mechanisms that offset some of the grain growth; the slight offset of the median DTM ratio above the blue line at high redshift is due to the fact that dust growth takes off very quickly. The transition from galaxies dominated by dust injected by stellar sources (mostly type II SNe) and that dominated by grain growth occurs at $z \sim 6$, as illustrated in Figure 2.4 which shows the fraction of galaxies in different DTM ratio bins.

The Rémy-Ruyer et al. (2015) data shown in Figure 2.3 combines two samples of local galaxies from the Herschel: Dwarf Galaxy Survey (DGS Madden et al. 2013, to study low-metallicity systems) and the Key Insights on Nearby Galaxies: a Far-Infrared Survey with Herschel (KINGFISH Kennicutt et al. 2011, mostly spiral galaxies along with several early-type and dwarf galaxies to include metal-rich galaxies). They use a semi-empirical dust SED model presented in Galliano, F. et al. (2011) to derive dust masses and estimate systematic errors of order 2–3. The DTM ratio predictions from our model show reasonable agreement with this data, although the dispersion in the model predictions is lower, and some of the highest observed DTM ratios are incompatible with the predictions of our model: the extent to which that is due to observational uncertainty is hard to assess.

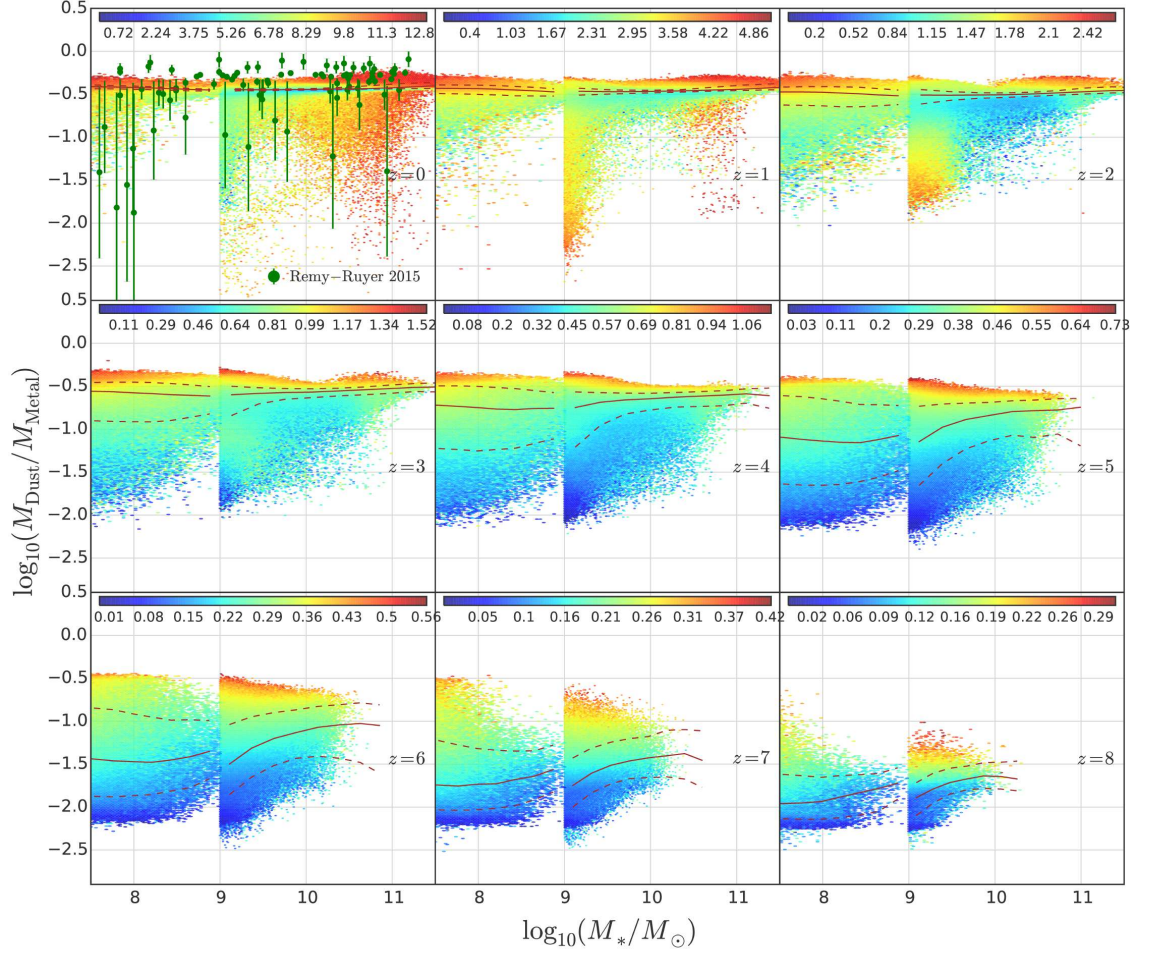


Figure 2.5: The dust-to-metal ratio as a function of stellar mass from $z = 0 - 8$, same as Figure 2.3, with galaxies coloured according to their mass-weighted stellar age in Gyrs. Green points show the observational constraints from Rémy-Ruyer et al. (2015).

The transition from the lower to the upper locus in Figures 2.3 & 2.4 is largely a function of the age of the galaxy – grain growth needs time to act (see also Appendix 2.A). This is shown clearly in Figure 2.5 which plots the same relation with galaxies coloured by their mass-weighted stellar age. Although the precise time taken for grain growth to saturate will depend upon the metallicity and initial dust content of the ISM, it takes of order 1 Gyr to do so. A study by Inoue (2003) has also shown that the evolutionary tracks in the metallicity – DTM ratio plane depends on the star-formation history.

At $z = 0$, the DTM ratio in some of the oldest, most massive galaxies has again begun to fall slightly and in some significantly – these are early types for which the molecular gas content of the cold ISM is low. We can therefore see that a galaxy’s DTM ratio depends strongly on it’s age, but also more weakly on it’s chemical enrichment, molecular gas consumption, and other factors relating to its evolutionary history.

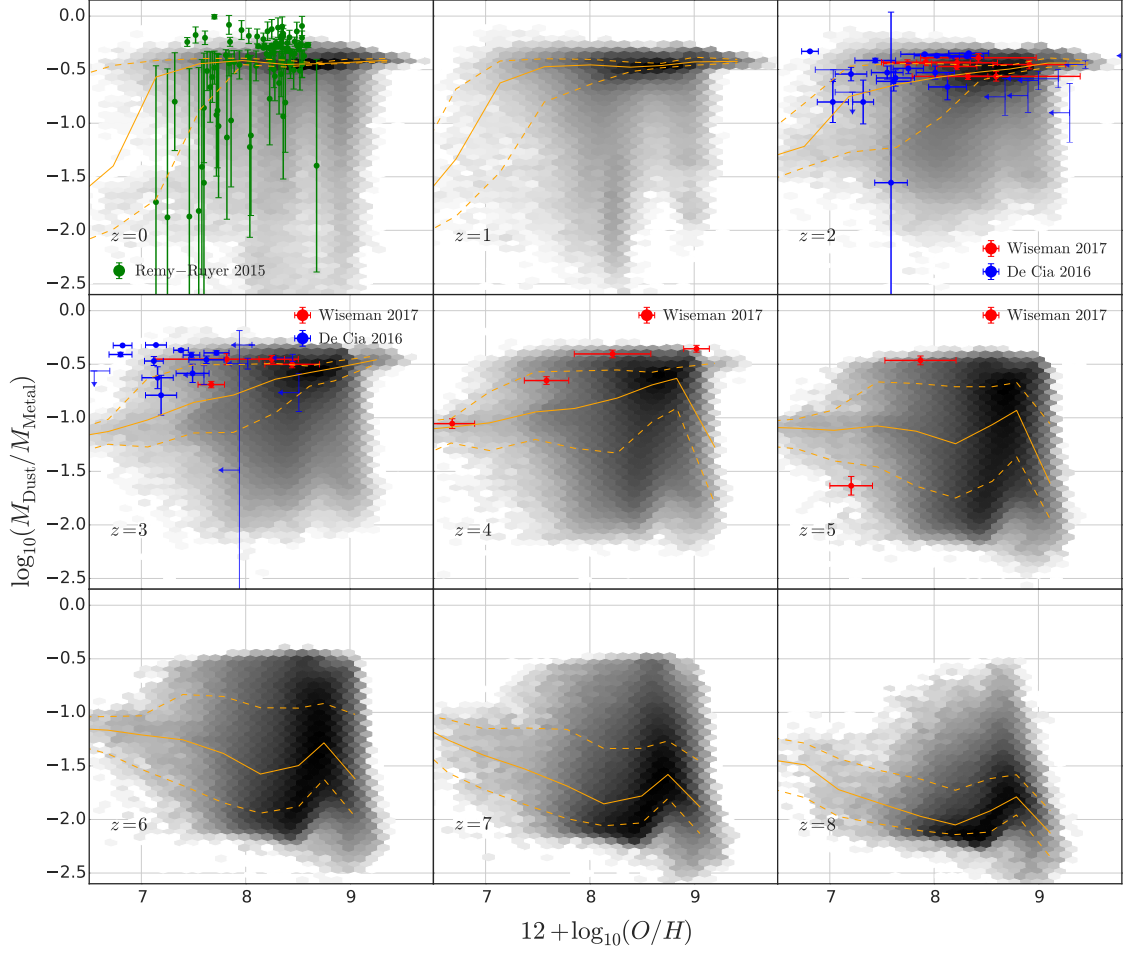


Figure 2.6: The dust-to-metal ratio as a function of their metallicity from $z = 0 - 8$. The orange line shows the median result from galaxies in our model, with the dashed lines denoting the 84 and 16 percentiles. Green blue and red points show the observational constraints from [Rémy-Ruyer et al. \(2015\)](#), [De Cia et al. \(2016\)](#) and [Wiseman et al. \(2017\)](#) respectively.

If we compare our results to [PSG17](#) (their Figure 6), our model galaxies do not exhibit any evolution of the DTM ratio with stellar mass as seen in their results at $z = 0$. But the scatter at $z = 0$, is negligible similar to [PSG17](#). At all redshifts their DTM ratio remains almost constant as well as exhibiting negligible scatter below $M_* = 10^{8.5} M_\odot$, while increasing rapidly afterwards due to their grain growth mechanism dominating the dust production. The cause of these differences are explained in Section 2.4.2.

2.4.1.2 DTM versus metallicity

Figure 2.6 shows the DTM ratio as a function of the gas-phase ISM oxygen abundance (i. e. the oxygen not locked up in dust). At $z = 0$, we again compare to observations from [Rémy-Ruyer et al. \(2015\)](#). We match the normalization of the observations for $12 + \log_{10}(O/H) > 8$ and also

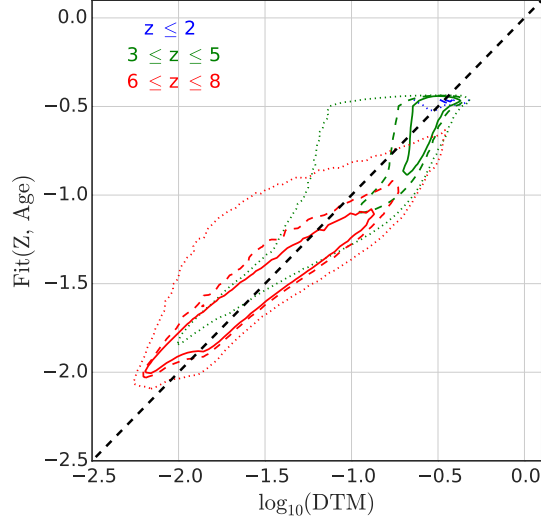


Figure 2.7: The DTM fitting function in Equation 2.15 is plotted against the DTM ratio from the model as contour enclosing different percentiles of the data. The blue, green and red contour lines represent model galaxies with $z \leq 2$, $3 \leq z \leq 5$ and $6 \leq z \leq 8$ respectively. Solid, dashed and dotted lines represent the 50, 68 and 95th percentile respectively. The dashed black line represents the one-to-one relation between the fitting function and the data.

some scatter down to low DTM ratios, noting that the low-DTM observational data tend to have the largest uncertainties. At higher redshifts, we show a good fit to the DTM ratios deduced by observations of gamma ray bursts (GRBs, from Wiseman et al., 2017) and damped lyman-alpha emitters (DLAs, from De Cia et al., 2016).

At $z \geq 6$ there appears a negative trend in the DTM-metallicity relation with increasing metallicity. This is due to the fact that at these high redshifts grain growth has not had sufficient time to enrich the cold ISM. This trend also emerges from the dust injection tables used in the model, since at these redshifts the DTM ratio follows the stellar injection modes of dust production. The same feature is seen in the model variants that are discussed in PSG17. The feature is absent in the fiducial model used in PSG17 due to their grain growth mechanism dominating the dust production (see Section 2.4.2).

2.4.1.3 DTM fitting function

As we have seen, the DTM ratio can vary by a large amount, depending upon the evolutionary history of a galaxy. It would be useful to be able to capture that behaviour with a suitable fitting function. Motivated by our conclusions earlier in this section, we posit the following functional form:

$$\text{DTM}_{\text{fit}} = \mathcal{D}_0 + (\mathcal{D}_1 - \mathcal{D}_0) \left[1 - \exp \left(-\alpha Z^\beta (\text{Age}/\tau)^\gamma \right) \right], \quad (2.15)$$

where \mathcal{D}_0 and \mathcal{D}_1 represent the initial type II SNe dust injection and the saturation value, respectively, Z is the metallicity of the interstellar medium, Age is the mass-weighted age of the stellar population, and $\tau = \tau_{\text{acc},0}/\mathcal{D}_0 Z$ is an estimate of the initial dust growth timescale after dust injection from type II supernovae but prior to the initiation of dust growth on grains.

Fixing the values of \mathcal{D}_0 and \mathcal{D}_1 by reference to Figure 2.3, the best fit values (using the Levenberg-Marquardt method implemented in the PYTHON package `scipy.optimize.curve_fit`) to the other parameters are:

$$\mathcal{D}_0 = 0.008,$$

$$\mathcal{D}_1 = 0.329,$$

$$\alpha = 0.017,$$

$$\beta = -1.337,$$

$$\gamma = 2.122.$$

The above fitting function is plotted against the DTM ratio in the model in Figure 2.7 for $z = 0 - 8$. The majority of galaxies lie close to the fit, well within about a factor of 2, although the full dispersion in DTM ratios is not quite captured. This then provides a good estimate of dust extinction should the metallicity and age of a galaxy be known, and offers a significant improvement upon the fixed DTM ratios often assumed in the literature (e.g. Wilkins et al., 2018). We show in the appendix that the same fitting function holds good for different choices of $\tau_{\text{acc},0}$.

2.4.2 Integrated dust production rates

The detailed dust model we have built includes several different dust production and destruction mechanisms that all contribute to the final dust properties of the galaxies in our model. Figure 2.8 shows the mean dust production (or destruction) rate densities as a function of redshift for galaxies in the $(480 h^{-1} \text{Mpc})^3$ MR simulation. The total dust destruction rate plotted includes destruction from supernovae, star formation, and reheating. We also plot the star formation rate density for comparison.

We can see that grain growth in molecular clouds dominates the production of dust over the redshift range $z = 0 - 8$, rapidly increasing towards its peak at $z = 2$. The destruction rate closely follows the dominant grain growth production rate, suggesting that any dust destroyed is rapidly recycled by grain growth. While at very early times type II supernovae dominate the

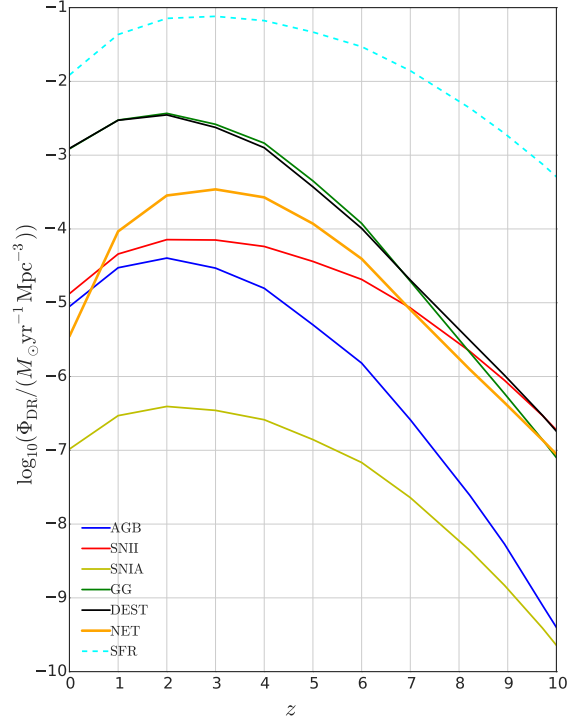


Figure 2.8: The production rate of dust through different mechanisms as a function of redshift for the MR run. Red, blue and yellow lines show the contribution from stellar sources of dust production, type II supernovae (SNII), AGB stars and type Ia supernovae (SNIA) respectively. The green line (largely obscured) shows the contribution from grain growth (GG) inside molecular clouds. The black line shows the dust destruction rate (DEST). The orange line shows the net total dust production rate (NET), taking into account all production and destruction. We also plot the star formation rate (SFR) density as a dashed cyan line for comparison.

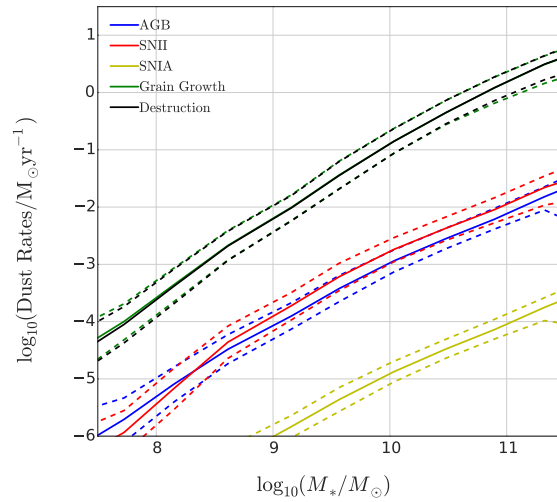


Figure 2.9: The production rate of different dust mechanisms as a function of stellar mass, shown for $z = 0$. Blue, red and yellow lines show the median contribution from stellar sources of dust production: AGB stars, type II supernovae and type Ia supernovae respectively, with the dashed lines denoting the 84 and 16 percentiles. The green line shows the contribution from grain growth inside molecular clouds. The black line shows the dust destruction rate.

production of dust. Thus, at the highest redshifts, the dispersion in the DTM ratio is small, with the dispersion increasing rapidly as grain growth takes over at $z < 8$.

If we look at the stellar contributions to the dust content, we see that type II supernovae are the dominant stellar production mechanism across the whole redshift range, peaking at $z \sim 2$, closely following the shape of the star formation rate as one would expect. Dust production (and metal enrichment) from Type Ia supernovae is shifted to slightly later times, due to the power-law delay-time distribution (DTD) we assume, which allows ~ 52 per cent of the supernovae to explode > 400 Myr after star formation (see [Yates et al. 2013](#)). Nonetheless, Type Ia supernovae never have a significant impact on the dust production rate. It is worthwhile to note that many other works also suggest that Type Ia supernovae are unlikely to be the major sources of ISM dust (e. g. [Nozawa et al., 2011](#)). Production by AGB stars is also negligible at early times, but rises at late times to rates approaching that of type IIs.

It is important to note that, although grain growth is the dominant dust formation mechanism at all redshifts below $z = 8$ when averaged over the whole galaxy population, the dust content of individual galaxies can vary enormously. At $z = 6$, for example, grain growth exceeds stellar dust injection by a factor of 6, but the spread in DTM ratios seen in [Figure 2.4](#) extends over more than a decade.

The variation of the dust production rates with stellar mass is shown in [Figure 2.9](#) for $z = 0$ for star forming galaxies (defined here as galaxies with a specific star formation rate, $\text{sSFR} > 1/3t_{\text{H}}(z)$, where $t_{\text{H}}(z)$ is the age of the Universe at redshift z). From this it is clear that there is very little dependence of dust growth and destruction upon galaxy mass. The same holds too at all other redshifts.

If we compare our dust production rates with the [PSG17](#) model (their Figures 8 & 10), this is bound to be different since the models differ in the grain growth implementation as well as the dust yield tables used for stellar production mechanisms. But the trends seen in both the models are similar in the sense that grain growth dominates over all the other production mechanisms at almost all redshifts from $z = 0 - 8$. In their model, the median grain growth rate is approximately 3 orders of magnitude higher than any stellar production mechanisms for high stellar mass ($> 10^{10} M_{\odot}$) at all redshifts. In our model, the dust production rate from SNI and grain growth is similar at $z \sim 8$. Also, we note that the production rates for each of the various sources (SNe-II, AGB stars, and grain growth) are similar between the two models at high mass, whereas they are three to four orders of magnitude greater at low mass in our model compared to [PSG17](#). These differences in the dust production rates from different processes are reflected in slight differences

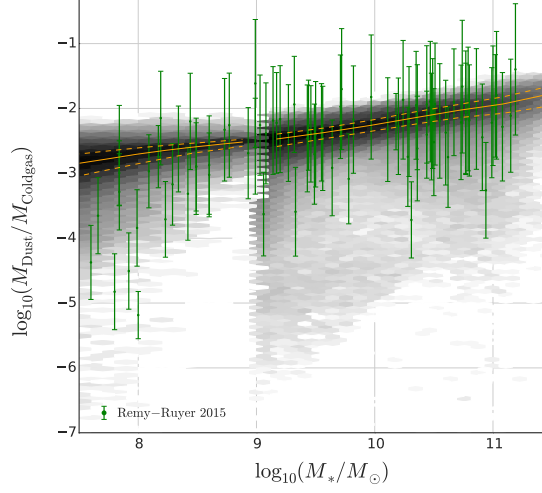


Figure 2.10: The dust-to-gas ratio as a function of stellar mass for $z = 0$. The orange line shows the median result from galaxies in our model, with the dashed lines denoting the 84 and 16 percentiles. Green points show the observational constraints from [R  my-Ruyer et al. \(2015\)](#).

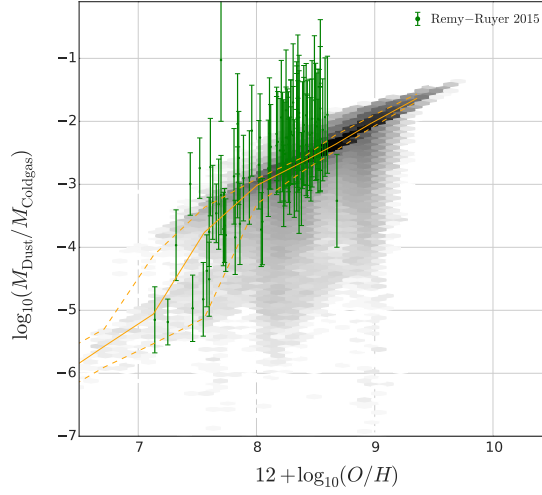


Figure 2.11: The dust-to-gas ratio as a function of gas phase metallicity for $z = 0$. The orange line shows the median result from galaxies in our model, with the dashed lines denoting the 84 and 16 percentiles. Green points show the observational constraints from [R  my-Ruyer et al. \(2015\)](#).

seen in our results for the dust content of galaxies, discussed in the next section.

2.5 Results: Dust content of galaxies

In this section we compare the predicted dust content of galaxies in our model to observations such as the dust-to-gas ratio, the stellar-mass – dust-mass relation and the dust mass function.

2.5.1 Dust-to-Gas (DTG) ratio

We compare the DTG ratio to two different properties, first, to see how the DTG ratio varies with stellar mass in Figure 2.10, and secondly how it varies with oxygen abundance, as seen in Figure 2.11. Because of the difficulty in obtaining observational data for comparison, we show only results for $z = 0$; at higher redshifts, the DTG ratio exhibits the same behaviour seen for the DTM ratio in Figure 2.3.

In Figure 2.10, we compare the DTG ratio of our model versus stellar mass against observations from Rémy-Ruyer et al. (2015). The median value of our model fits the observations well, particularly above stellar masses of $10^8 M_\odot$. Below this value, there may be a downturn in the DTG ratio in the data, that we do not see. Figure 2.11 shows the same data plotted as a function of oxygen abundance and here we see that the low DTG ratios are associated with low metal abundance, and that the observations and the model overlap quite well. The reason for the discrepancy seen at low masses in Figure 2.10 is therefore due to the fact that our low-mass galaxies mostly have higher oxygen abundance than those in the Rémy-Ruyer et al. (2015) sample.

The PSG17 (their Figure 4 and 3) as well as Hou et al. (2019) (their Figure 4a and 4b) model exhibits a similar trend to our predictions when the DTG ratio is plotted as a function of stellar mass and metallicity respectively. But at all redshifts both the models exhibit a steeper slope, such that their lower mass model galaxies have lower DTG ratios. In case of McKinnon et al. (2017), the DTG ratio shows a flat trend with metallicity (their Figure 8) for $12 + \log(\text{O}/\text{H}) > 8$ while showing a positive correlation below that.

2.5.2 Dust versus stellar mass

The dust mass versus stellar mass relation is shown in Figure 2.12. The evolution in dust masses mimics that shown in Figure 2.3 for the DTM ratio. At $z = 0$ most of the galaxies have saturated dust growth on grains. This persists up to $z = 4$, after which there is a gradual transition down to the levels expected for dust injection from stellar sources.

The stellar-dust mass parameter space is one where we have observational constraints across a very large range of redshifts. The coloured points in Figure 2.12 represent observations from a number of different studies (DustPedia collaboration Davies et al., 2017; Ciesla et al., 2014; da Cunha et al., 2015; Mancini et al., 2015; Rémy-Ruyer et al., 2015; Santini et al., 2014). The DustPedia data combine the Herschel/Planck observations with that from other sources of data, and provide observations at numerous wavelengths across the spectral energy distribution. The

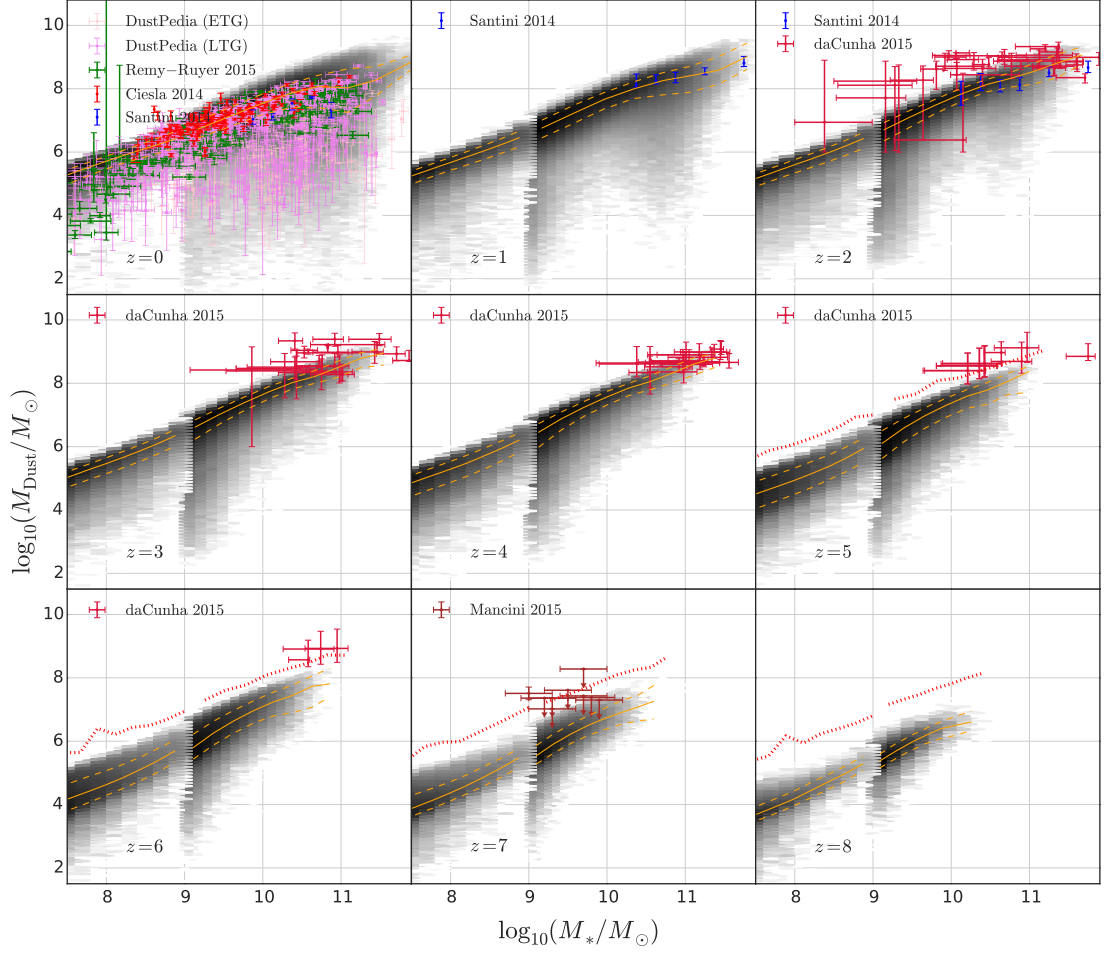


Figure 2.12: The stellar-dust mass relation for redshifts $z = 0 - 8$. The orange line shows the median result from galaxies in our model, with the dashed lines denoting the 84 and 16 percentiles. Pink, violet, red, green, blue, crimson and brown points show the observational constraints from the DustPedia archive (see [Davies et al., 2017](#), separated into ETGs and LTGs), [Ciesla et al. \(2014\)](#), [Rémy-Ruyer et al. \(2015\)](#), [Santini et al. \(2014\)](#), [da Cunha et al. \(2015\)](#) and [Mancini et al. \(2015\)](#) respectively. At redshifts 5 and above, the dotted red line shows the maximal dust content that could be predicted by our model, assuming saturated grain growth and no dust destruction.

dust masses are fitted using CIGALE assuming either the dust model from [Draine et al. \(2014\)](#) or their own called THEMIS (see [Davies et al. 2017](#)). We use the dust masses fitted by the former model, since the latter has a lower normalisation at $z = 0$ compared to our dust masses. The [Ciesla et al. \(2014\)](#) data uses the *Herschel* Reference Survey ([Boselli et al., 2010](#)), where the dust masses are obtained using the SED templates described in [Draine & Li \(2007\)](#). [da Cunha et al. \(2015\)](#) derives dust masses from a sample of sub-mm galaxies in the ALMA LESS survey using the SED fitting techniques described in [da Cunha et al. \(2008\)](#). Some of the galaxies in the sample only have photometric redshifts and thus the redshift is kept as a free parameter in their fitting technique. [Mancini et al. \(2015\)](#) uses ALMA and PdBI observations with upper limits on the dust continuum emission. They derive the stellar masses using the mean relation between the UV

magnitude and the dust mass assuming $T_d = 35\text{K}$ and $\beta = 1.5$. [Santini et al. \(2014\)](#) uses galaxies in the GOODS-S and GOODS-N field as well as the COSMOS field which have FIR observations carried out using *Herschel*. They also use the SED templates of [Draine & Li \(2007\)](#) as a description for their dust masses.

The first thing to note is that there is a significant offset in normalisation between the different observational data sets at $z = 0$. Thus we see that, while the median dust content predicted by our model is consistent with the LTGs from DustPedia and [Ciesla et al. \(2014\)](#) data, it lies well above that of [Rémy-Ruyer et al. \(2015\)](#) and [Santini et al. \(2014\)](#). This reflects the different observational biases and systematic uncertainties in the estimation of dust content. For example, the [Rémy-Ruyer et al. \(2015\)](#) sample contains some massive AGN-host galaxies which are presumably older and have low gas fractions, leading to smaller dust masses. Also a part of their sample (DGS, [Madden et al., 2013](#)) was chosen to study low-metallicity environments and hence exhibit smaller dust masses.

Although the median dust level is acceptable, it would appear that we have many galaxies, particularly at masses above $\sim 10^{10} M_\odot$, whose dust content is significantly higher than those seen in the observational samples considered here. This could come about in one of three ways: too much cold gas; too high a metallicity in the cold gas; too high a dust-to-metal (DTM) ratio. The cold gas content of galaxies in the [HWT15](#) model was considered in [Martindale et al. \(2017\)](#) and while the H I mass function was in good agreement with the observations, the gas-to-stellar mass ratio is, if anything, slightly too low (although the selection functions for the H I surveys are hard to reproduce). Similarly, [Yates et al. \(2013\)](#) showed that the oxygen abundance of cold gas in our model is in good agreement with observations from SDSS. Finally, Section 2.4.1 of this paper shows that the DTM ratio is in good agreement with that of [Rémy-Ruyer et al. \(2015\)](#). It is thus slightly perplexing that we seem to have these galaxies with excessive dust. We note that in our model, we have ignored possible dust destruction due to the effects of cosmic rays, photoevaporation or AGN activity that start to play a major role in high mass galaxies.

There is also a significant spread in observed dust masses to lower values at high stellar masses at $z = 0$ due to the presence of elliptical early-type galaxies (ETGs) with low molecular gas content. We predict many such galaxies in our model (see also Figure 2.5) but in a lower proportion than in the DustPedia data set – it is unclear to what extent this is an observational selection effect.

At higher redshifts, up to $z = 4$, the upper locus of our dust masses lies, if anything, slightly below the observations, and at $z = 5, 6$ and 7 it is well below. We note, however, that almost

all of the Mancini data are upper limits, and that the [da Cunha et al. \(2015\)](#) data are *ALMA* observations of sub-mm galaxies which are some of the brightest star-forming galaxies at that particular redshift, hence a population biased towards more dust-rich systems. The dotted red lines in Figure 2.12 show the saturation value (as discussed in Section 2.4.1). To reproduce any observations lying above this would require either a higher cold gas content, or a higher metallicity (i.e. earlier enrichment), or too low a dust destruction rate in the semi-analytic model. It is worthwhile to note that the dust destruction efficiency adopted in this study is based on calculations for multiphase ISM in the solar environment, hence one could imagine the ISM having different properties at $z > 5$, thus also changing the dust destruction rates.

[PSG17](#) also found mixed success in matching observations of the stellar mass – dust mass relation (their Figure 2) in both local and high-redshift galaxies. At $z = 0$, their median relation lies below the observations of [Ciesla et al. \(2014\)](#), but follows the trend seen by [Rémy-Ruyer et al. \(2015\)](#) at low mass, where they reproduce a steep stellar mass – dust mass relation. This is chiefly due to the longer accretion timescales they assume at low molecular gas densities, which can reach around 1 Gyr (see their Figure 1), compared to values closer to 10 Myr for this work (see Figure 2.14). They have galaxy masses up to $3 \times 10^{11} M_{\odot}$ at all redshifts up to $z = 9$, finding a median dust-to-stellar mass relation with a steeper slope than our results, thus providing a better match to the high redshift observations than we do. We note that these differences in our results are driven by the strong molecular-gas dependence in their empirical τ_{acc} prescription, which is in turn driven by the enhanced star-formation efficiency they assume at $\Sigma_{\text{H}_2, \text{crit}} > 70 M_{\odot}/\text{pc}^{-2}$ (their Equation 1); our model assumes much smaller variations in the properties of molecular clouds in galaxies of different surface densities.

2.5.3 Dust mass function

Figure 2.13 shows the dust mass function at $z = 0$. The red line shows the results of the Millennium-II simulation, and the black line the Millennium simulation. We compare with observations from [Dunne et al. \(2003\)](#); [Vlahakis et al. \(2005\)](#); [Eales et al. \(2009\)](#); [Dunne et al. \(2011\)](#); [Clemens et al. \(2013\)](#). [Dunne et al. \(2003\)](#) obtained data for the local and high-redshift dust masses using the *SCUBA* (*Submillimetre Common-User Bolometer Array*) Local Universe Galaxy Survey (SLUGS) and for high-redshift ($z = 2.5$) submillimetre data from the deep SCUBA submillimetre surveys. For the local objects a dust temperature of 20 K was used while for the high-redshift sample dust temperature of 25 K was used, thus assuming an increase in the average dust temperature of galaxies. [Vlahakis et al. \(2005\)](#) derived the local sub-mm luminosity and dust mass functions using SLUGS and the *IRAS* Point Source Catalog Redshift Survey (PSCz).

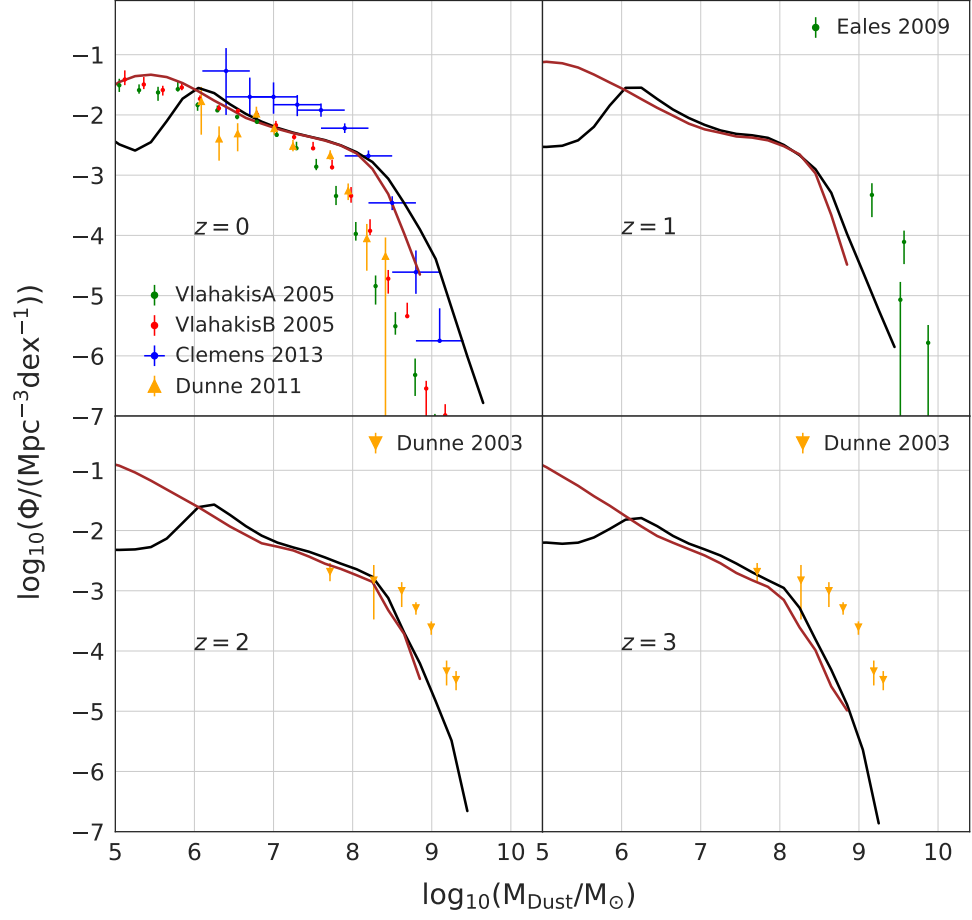


Figure 2.13: The Dust Mass Function (DMF) for redshifts $z \in [0, 3]$. The black line shows the prediction of our model using the underlying dark matter Millennium simulation, and the red line for Millennium-II. Observations are taken from Dunne et al. (2011), Vlahakis et al. (2005) and Clemens et al. (2013) at $z = 0$, Eales et al. (2009) at $z = 1$ and Dunne et al. (2003) for $z = 2.5$.

They fit two component grey bodies to their SEDs with emissivity index $\beta = 2$ and dust temperature in the range 17-24 K. The ‘A’ sample determines dust masses using a dust temperature obtained from isothermal SED fitting, and the ‘B’ dust mass function has been calculated using a dust temperature of 20 K. Eales et al. (2009) uses data obtained from the *Balloon-borne Large Aperture Submillimeter Telescope (BLAST)*, using the greybody relation assuming a dust temperature of 20 K. Dunne et al. (2011) using the *Herschel*-Astrophysical Terahertz Large Area Survey (*Herschel*-ATLAS) calculated dust masses fitting a single-temperature grey-body model for the spectral energy distribution with $\beta = 1.5 - 2.0$ and dust temperature in the range 10 – 50 K. Clemens et al. (2013) combined *Herschel* data with *Wide-field Infrared Survey Explorer (WISE)*, *Spitzer* and *Infrared Astronomical Satellite (IRAS)* observations to investigate the properties of a flux-limited sample of local star-forming galaxies. They fit their SEDs with modified blackbody spectra using $\beta \simeq 2$ and dust temperatures in the range 10-25 K.

We find that the model provides a good fit to the Vlahakis et al. (2005) and Dunne et al. (2011)

observations at low and intermediate dust masses, but under-predicts the number density when compared with [Clemens et al. \(2013\)](#) at the same mass range. The knee of the mass function is at a lower mass in [Dunne et al. \(2011\)](#); [Vlahakis et al. \(2005\)](#) compared to our model output, while in [Clemens et al. \(2013\)](#) it roughly coincides with our model. At the high mass end, our predicted number densities are higher than both the observational data sets. This result is consistent with that of the previous section, that we over-predict the dust content of many massive galaxies at $z = 0$ in our model. On comparing our model predictions to the [Dunne et al. \(2003\)](#); [Eales et al. \(2009\)](#) data for $z \geq 1$, we instead appear to slightly under-predict the dust mass function at high masses. It is worthwhile to note that this is a general feature seen in other models of galaxy formation tracking dust growth (e. g. [McKinnon et al. 2016](#), [PSG17](#)).

2.6 Conclusions

We have run a modified version of the L-GALAXIES semi-analytic model which includes a prescription of dust modelling on the full Millennium and Millennium-II trees. By combining both the Millennium simulations we are able to make use of both the higher volume in order to find rarer objects, but also the higher mass resolution of Millennium-II to probe lower mass galaxies. Our conclusions are as follows:

1. Our grain growth model follows that of previous work, as described in [Popping et al. \(2017a\)](#), but following separately the dust content in molecular clouds and the inter-cloud medium. We find that, in regimes where $\tau_{\text{exch}} \gg \tau_{\text{acc}}$ as well as for low values of μ , this can have a significant impact upon the dust growth rate (Figure 2.1).
2. The dust-to-metal (DTM) ratio (Figure 2.3) shows an evolution from low to high ratios, the former corresponding to dust injection from type II supernovae, and the other to maximal, saturated dust production occurring via dust growth on grains. The latter dominates at redshifts below $z \approx 4$. A significantly populated transition region is seen at $z = 6$.
3. By colouring with age (Figure 2.5) we show that this is the primary driver of the movement from low to high DTM ratio.
4. When plotted as a function of gas-phase metallicity, we find a reasonable fit to the observations at all redshifts (Figure 2.6).
5. We present a fitting relation for the DTM ratio, dependent on the metallicity and mass-weighted age of the galaxy stellar population. That provides a good fit to the model at

both low and high redshift, but with some scatter at intermediate redshifts due to the varied growth histories of galaxies (Equation 2.15 and Figure 2.7).

6. Grain growth is the dominant dust production mechanism at all redshifts below $z = 8$ (Figure 2.8). Dust destruction rate closely follows the grain growth production rate, suggesting prompt recycling of any dust content. We note, however, that Figure 2.3 shows that by $z = 6$ only half of galaxies lie on the upper locus of DTM ratio. Thus the detailed history of galaxy formation is important for determining the dust content of any individual galaxy.
7. The dust growth rates show little dependence on galaxy mass (Figure 2.9).
8. We find a good fit to the shape and normalisation of the dust-to-gas ratio at $z = 0$ when plotted as a function of both stellar mass (Figure 2.10) and oxygen abundance (Figure 2.11).
9. We find a reasonable fit to the shape and normalisation of the observations in the stellar-dust mass plot (Figure 2.12) over a wide range of redshifts, $z = 0 - 4$. We have an excess of very dusty, massive galaxies at $z = 0$, perhaps due to a lack of destruction mechanisms. We fail to predict the dustiest galaxies at $z > 5$, which hints that our dust growth rate may be too slow, or the destruction rate too high; however, we note that the interpretation of the observations are very uncertain at these redshifts.
10. There is a good agreement between the predicted $z = 0$ dust mass function at the intermediate and low dust masses with observations; however we over predict the number density of galaxies at the highest dust masses (Figure 2.13). This again suggests that we may have too much dust in the most massive galaxies.

The model that we have presented here is deficient in at least 2 respects. Firstly, it assumes that dust is instantly destroyed in the hot (coronal) phase of the interstellar medium. Secondly, we ignore the effect of dust on the physics of galaxy formation: the formation of molecules on grains, and the coupling to radiative feedback, for example. This will be investigated in future work.

It seems evident from our work that, at sufficiently high redshift, there will be a transition from high (saturated dust growth) to much lower (primarily type II supernovae) dust-to-metal ratios. The precise redshift at which this happens depends upon uncertain grain growth and destruction time-scales. Nonetheless, it is important to appreciate that there will be a wide variety of DTM ratios in galaxies at high redshift. The situation will become much clearer over the next

few years with deep extragalactic surveys such as those proposed by *Euclid*, *Roman* and *JWST* and follow-up with ground-based observations from facilities such as *ALMA*.

2.A Accretion Timescale

Here we will discuss how the accretion timescale varies with redshift as well the impact of choosing a different $\tau_{\text{acc},0}$ on our dust model.

Figure 2.14 shows the distribution of τ_{acc} plotted against stellar mass for $z = 0 - 8$. The age of the Universe at each redshift is also plotted for comparison. The median value of τ_{acc} moves towards lower values as we move to lower redshifts due to the increase in the DTG ratio (see equation 2.11). Note that there are a lot of galaxies at high redshift ($z \geq 6$) that have τ_{acc} values similar to the age of the Universe at that particular redshift – this is also the reason for very low values of DTG or DTM ratio, with comparable or higher values of the stellar production rate compared to grain growth. As we move towards lower redshift, most τ_{acc} values start to dip beneath the age of the Universe and at $z \leq 2$ the median values are 3 to 4 orders of magnitude less than the age of the Universe. Thus the choice of $\tau_{\text{acc},0}$ has a negligible effect at low redshifts but can be quite significant in determining the galaxy dust mass at high redshifts.

To see the effect of modifying the value of $\tau_{\text{acc},0}$ on the galaxy dust mass we consider values ranging from $5 \times 10^3 - 10^6$ yr. The median dust-stellar mass relation for $z = 0, 2, 5$ and 6 with these accretion timescales are shown in Figure 2.15. The dust-stellar mass relation at $z = 0$ is not drastically affected by changes in $\tau_{\text{acc},0}$, except for $\tau_{\text{acc},0} = 10^6$ yr where the median is about 0.5 dex lower than the other median values at intermediate stellar masses – this is because the dust growth timescale becomes comparable to the destruction timescale. Similarly, at $z = 2$ the DTM ratio for $\tau_{\text{acc},0} = 10^6$ yr has decreased by more than an order of magnitude. At $z = 5$ and 6 the changes are more visible with a spread in DTM ratios becoming apparent as $\tau_{\text{acc},0}$ is varied. The main point to take away from this is that grain growth requires time to act, and that timescale depends on the value of $\tau_{\text{acc},0}$.

We also compare how our fitting function, Equation 2.15 performs for different values of $\tau_{\text{acc},0}$ in Figure 2.16. For this we ran our model with $\tau_{\text{acc},0}$ values of 10^4 , 10^5 and 10^6 yr, and obtain the expected DTM ratio using the corresponding $\tau_{\text{acc},0}$ values in Equation 2.15. We see that the fit does a good job for $\tau_{\text{acc},0} = 10^4$ yr where we expect DTM ratios to be near saturation, while for the higher $\tau_{\text{acc},0}$ values we see considerable scatter in the fit. This scatter for $\tau_{\text{acc},0} = 10^6$ yr directly follows from our previous discussion of the grain growth timescales. This has led to a bimodal distribution at $z \leq 2$, with the grain growth dominated population at the top and the

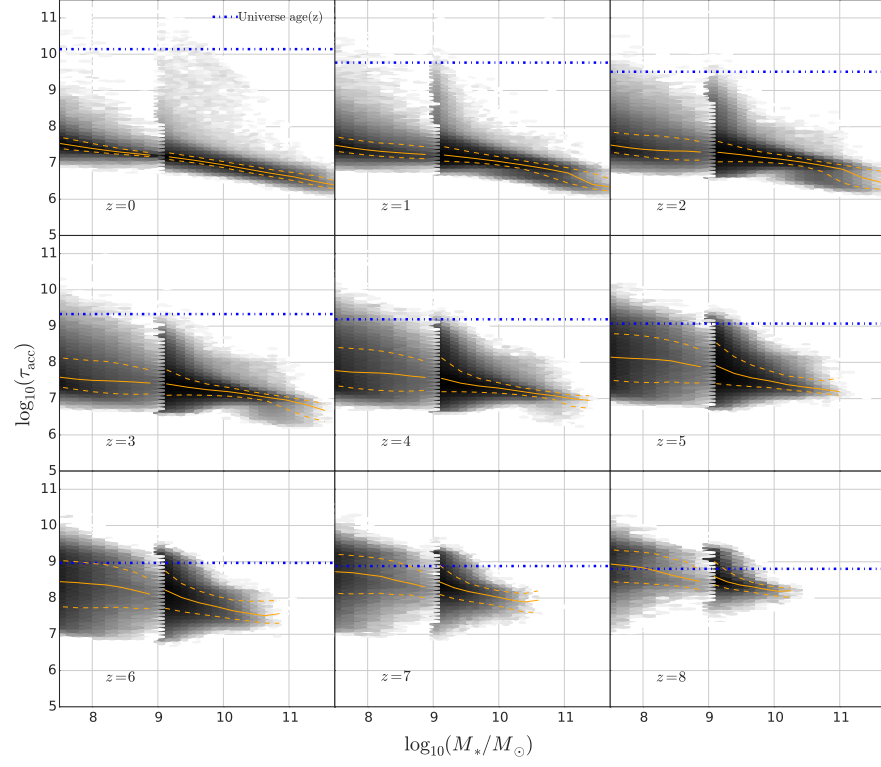


Figure 2.14: The accretion timescale-stellar mass relation for redshifts $z = 0 - 8$. The orange lines show the median result from galaxies in our model and the 1-sigma scatter. The age of the Universe at that particular redshift is shown as the dot-dashed blue line.

stellar injection dominated ones at the bottom. The sharp cut-off in the bottom population is an artifact of our fitting function, as it can not have values less than \mathcal{D}_0 .

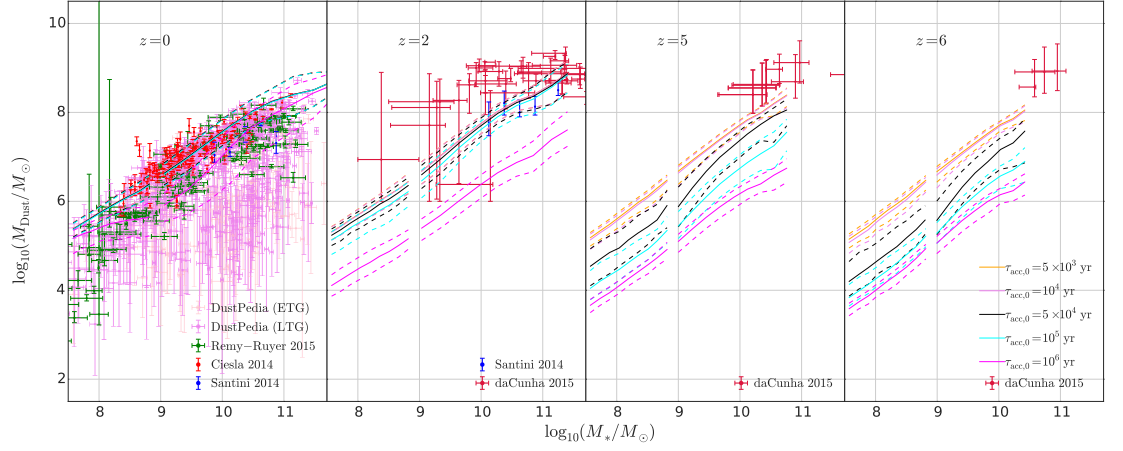


Figure 2.15: The median dust-stellar relation in our model for different values of $\tau_{\text{acc},0}$ at $z = 0, 2, 5$ and 6 . The solid line shows the median relation while the dashed lines denotes the 84 and 16 percentiles. The observational constraints from the DustPedia archive (see [Davies et al., 2017](#), separated into ETGs and LTGs), [Ciesla et al. \(2014\)](#), [R  my-Ruyer et al. \(2015\)](#), [Santini et al. \(2014\)](#) and [da Cunha et al. \(2015\)](#) respectively have been plotted for comparison.

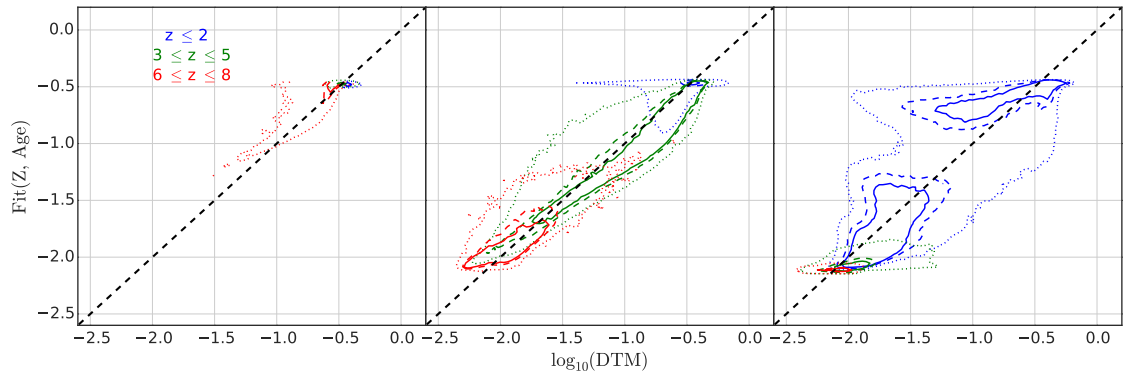


Figure 2.16: The DTM fitting function in Equation 2.15 is plotted against the DTM ratio from the model, similar to Figure 2.7. From left to right DTM values generated by running our model for $\tau_{\text{acc},0}$ values of 10^4 , 10^5 and 10^6 yr respectively are plotted against the fit function.

INTRODUCTION TO FIRST LIGHT AND REIONISATION EPOCH SIMULATIONS (FLARES)

3

This chapter serves as an introduction to the suite of simulations collectively called the First Light And Reionisation Epoch Simulations (FLARES). All the sections detailed here are from [Lovell et al. \(2021a\)](#), with the introduction of this chapter containing ideas and narrative from that work's introduction as well as from [Vijayan et al. \(2021\)](#) (the next chapter, introducing the photometric properties), to avoid repetition of themes. Also removed is the section on the star forming sequence, which I have not contributed to.

In FLARES, we resimulate a range of overdensities during the Epoch of Reionisation (EoR) in order to build composite distribution functions, as well as explore the environmental dependence of galaxy formation and evolution during this critical period of galaxy assembly. The regions are selected from a large $(3.2 \text{ cGpc})^3$ parent volume, based on their overdensity within a sphere of radius $14 \text{ h}^{-1} \text{ cMpc}$. We then resimulate with full hydrodynamics, and employ a novel weighting scheme that allows the construction of composite distribution functions that are representative of the full parent volume. This significantly extends the dynamic range compared to smaller volume periodic simulations. This chapter presents the galaxy stellar mass function (GSMF) and the star formation rate distribution function (SFRF) predicted by FLARES, and compare to a number of observational and model constraints. Also analysed is the environmental dependence over an unprecedented range of overdensity. Both the GSMF and the SFRF exhibit a clear double-Schechter form, up to the highest redshifts ($z = 10$). The increased dynamic range probed by FLARES will allow us to make predictions for a number of large area surveys that will probe the EoR in coming years, carried out on new observatories such as *Roman* and *Euclid*.

3.1 Introduction

The past few decades have seen tremendous growth in the understanding of galaxy formation and evolution in the first billion years of the Universe after the Big Bang. The first stars and galaxies formed within the first few million years after the big bang were the first sources of ionising photons in the Universe, ushering in the Epoch of Reionisation (EoR) by ionising hydrogen (e. g. Wilkins et al., 2011a; Bouwens et al., 2012; Robertson et al., 2013, 2015; Dayal & Ferrara, 2018).

Thanks chiefly to the efforts of the *Hubble Space Telescope* (HST, e. g. Beckwith et al., 2006; Bouwens et al., 2008; Labbé et al., 2010; Robertson et al., 2010; Wilkins et al., 2010; Bouwens et al., 2014; McLeod et al., 2015; Bowler et al., 2017; Kawamata et al., 2018) and the *Visible and Infrared Survey Telescope for Astronomy* (VISTA, e. g. Bowler et al., 2014; Stefanon et al., 2019; Bowler et al., 2020) more than a thousand galaxies have now been identified at $z > 5$ with a handful of candidates even identified at $z > 10$ (e. g. Oesch et al., 2016; Bouwens et al., 2019). These efforts have also been complemented by *Spitzer* providing rest-frame optical photometry (e. g. Ashby et al., 2013; Roberts-Borsani et al., 2016; Bridge et al., 2019) and the *Atacama Large Millimeter/submillimeter Array* (ALMA, e. g. Smit et al., 2018; Carniani et al., 2018; Hashimoto et al., 2019) providing rest-frame far-IR and sub-mm photometry and spectroscopy.

With upcoming facilities like the *James Webb Space Telescope*, *Euclid*, and the *Nancy Grace Roman Space Telescope* that can comprehensively study galaxies in the EoR, it is timely to model and predict the properties of these high redshift systems. The *Webb Telescope* will be able to provide better sensitivity and spatial resolution in the near and mid-infrared, providing rest-frame UV-optical imaging and spectroscopy. *Euclid*, and *Roman Space Telescope* can do deep and wide surveys adding better statistics to the bright end. The combined efforts of both these observatories can thus provide effective constraints on the bright and rare galaxies in the early Universe. These next generation of surveys would be the test beds to further the theory of galaxy formation and evolution.

In order to complement this upcoming phase of incredible wealth of data, theoretical works on galaxy formation and evolution should be built to explore these galaxy populations. Many theoretical works on simulations of galaxy evolution have already been used to study the population of galaxies and their properties in the EoR (e. g. Mason et al., 2015; Wilkins et al., 2017; Ceverino et al., 2017; Ma et al., 2018; Finlator et al., 2018; Yung et al., 2019a; Wu et al., 2020). There are various intrinsic physical properties of galaxies, like stellar mass and star formation rate, that are available directly from simulations, which can be compared to that of observed galaxies. These all involve some modelling assumptions based on the star formation history or metallicity of the

observed galaxies, which are hard to derive with limited available data on the galaxy at these high redshifts. Another approach is to make predictions from simulations to compare to galaxy observables that suffer from comparatively less modelling biases such as luminosities and line equivalent widths, thus providing insights into the physical processes that take place in these galaxies.

A goal of such numerical galaxy evolution studies is to model a representative population of galaxies, resolving all of the relevant physics at the required scales, in order to provide a test bed for the study and interpretation of observed galaxies (Benson, 2010). In order to achieve this it is necessary to simulate large volumes (in order to sample a representative volume of the Universe) at high resolution (*e.g.* spatial, mass, time; in order to resolve the internal physical processes within individual galaxies) and with all of the key physics included (such as full hydrodynamics, magnetic fields, *etc.*). Unfortunately this is not computationally feasible; compromises must be made with volume, resolution or choice of physics, depending on the scientific questions posed (for a review, see Somerville & Davé, 2015).

Predictions for these upcoming surveys have so far typically been made using phenomenological models. One such class of methods are Semi-Analytic Models (SAMs), run on halo merger trees extracted from dark matter-only simulations (for a review, see Baugh, 2006). Due to their efficiency they can be applied to large cosmological volumes, and used to probe distribution functions of intrinsic properties and observables over a large dynamic range. A number of these models have been tested during the EoR (Clay et al., 2015; Somerville et al., 2015; Poole et al., 2016; Rodrigues et al., 2017; Lagos et al., 2019; Yung et al., 2019a; Dayal et al., 2020; Hutter et al., 2021). Mock observables can also be produced and directly compared with observed luminosity functions (Lacey et al., 2016; Yung et al., 2019a). Such models can be run relatively quickly, allowing parameter estimation through Monte Carlo approaches (Henriques et al., 2015, 2020), a powerful means of exploring large degenerate parameter spaces. With each generation of SAMs, there are more detailed physical models being incorporated in them. However they treat galaxies as unresolved objects, modelling various components of galaxy evolution with their integrated properties. Hence, they do not self-consistently evolve various interactions such as mergers, feedback events or outflows, requiring additional steps and approximations to retrieve observables.

In contrast, hydrodynamical simulations of galaxy formation model in greater detail the evolution of dark matter, gas, stars and black holes, allowing for a more detailed exploration of galaxy structure and observed properties. Many state of the art periodic cosmological volumes like MASSIVEBLACK (Matteo et al., 2012), ILLUSTRIS (Vogelsberger et al., 2014a,b; Genel et al., 2014;

Sijacki et al., 2015), Horizon-AGN (Dubois et al., 2014), MASSIVEBLACK-II (Khandai et al., 2015), EAGLE (Schaye et al., 2015; Crain et al., 2015), MUFASA (Davé et al., 2016), ILLUSTRIS-TNG (Naiman et al., 2018; Nelson et al., 2018; Marinacci et al., 2018; Springel et al., 2018; Pillepich et al., 2018), Romulus (Tremmel et al., 2017), SIMBA (Davé et al., 2019), etc have been undertaken independently down to $z \sim 0$ with mass resolutions of order $10^6 M_\odot$, sufficiently high to resolve the internal structure of galaxies. However, their volumes are too small to replicate many of the current observations of bright massive galaxies, which are born in rare overdensities in the EoR. The enormous computational time to run such large periodic volumes have been a major roadblock from exploring large dynamic ranges with better resolution.

Most existing periodic hydrodynamic simulations during the EoR are not able to achieve the large dynamic ranges accessible by SAMs. This is illustrated in Figure 3.1, which shows where a number of existing simulations lie on a plane of simulated volume against hydrodynamic element mass. There is a strong negative correlation, with some outliers. The BLUE TIDES simulation (Feng et al., 2016), based on the Massive Black suite of simulations (Matteo et al., 2012; Khandai et al., 2015), was performed within a $(500 / h \text{ cMpc})^3$ periodic box, ~ 125 times as massive as the fiducial EAGLE reference volume, whilst at a similar resolution. They make predictions for a number of intrinsic and observational properties during the EoR (e.g. Waters et al., 2016; Di Matteo et al., 2017; Wilkins et al., 2016b; Wilkins et al., 2016c, 2017, 2018, 2020). Unfortunately, due to the increased computational cost it has only been run down to $z = 7$, and the model cannot therefore be tested against low redshift observables. Other simulations have taken a different approach, instead simulating smaller volumes at much higher resolution, allowing them to investigate the effect of a number of physical processes in greater detail (O’Shea et al., 2015; Rosdahl et al., 2018; Jaacks et al., 2019). However, these must similarly be stopped at intermediate redshifts due to the higher computational expense.

A successful approach to tackle this limitation has been the use of zoom simulations, whose regions are drawn from less expensive, low-resolution dark matter only simulations, whose box lengths can be in the gigaparsecs. These can be run at higher resolution with additional physics, by generating the initial conditions of the required patch of volume. This approach preserves the large-scale power and the long-range tidal forces by simulating the matter outside the volume of interest at a much lower resolution. For instance, this technique has been successfully employed to re-simulate cluster environments (similar to the works of Bonafede et al., 2011; Planelles et al., 2014; Pike et al., 2014, etc) in the C-EAGLE simulations (Barnes et al., 2017b; Bahé et al., 2017), whose regions were selected from a parent dark matter only simulation box of side length 3.2 cGpc (Barnes et al., 2017a). The simulations used the EAGLE physics model, allowing the model

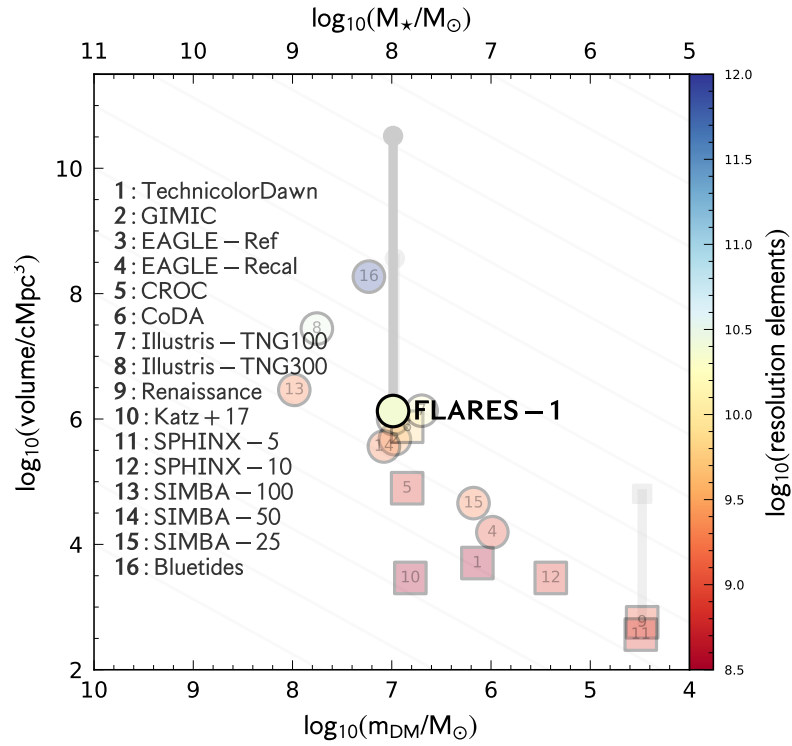


Figure 3.1: Dark matter element resolution against simulated volume. The colour of individual points describes the approximate number of resolution elements (dark matter + baryonic gas, excluding stars). We show the following simulation projects: Technicolor Dawn (Finlator et al., 2018), GIMIC (Crain et al., 2009), EAGLE (Schaye et al., 2015; Crain et al., 2015), CROC (Gnedin, 2014), CoDa (Ocvirk et al., 2016), Illustris (Vogelsberger et al., 2014a), Renaissance (Barrow et al., 2017), the Katz et al. (2017) simulations, SPHINX (Rosdahl et al., 2018), and BLUE TIDES (Feng et al., 2016). We also show FLARES with the total resimulated high-resolution volume, as well as a vertical line showing the representative volume, given by that of the parent box. There is a strong negative correlation for periodic volumes between the volume that can be simulated and the resolution that can be achieved. The resimulation approach, with appropriate weighting, allows us to extend the volume axis significantly.

to be used in cluster environments without the need to simulate large periodic boxes. There have also been high resolution zoom simulations that have probed the galaxy properties in the EoR like the stellar mass function or the luminosity function (e. g. Ceverino et al., 2017; Ma et al., 2018) as well as the Lyman- α /Lyman-continuum studies (e. g. Katz et al., 2018) or line emissions (e. g. Pallottini et al., 2019). However they have not necessarily extended the dynamic range that will be probed by the next generation surveys.

The zoom technique can also be applied to get representative samples of the Universe. An example of this, was the GIMIC simulations (Crain et al., 2009), which sampled 5 regions of various overdensities from the dark matter only Millennium simulation (Springel et al., 2005a) at $z = 1.5$. These regions were then re-simulated at a higher resolution with full hydrodynamics. In this case one can produce composite distribution functions by combining the regions using appropriate weights based on their overdensity. This allows for the exploration of the environ-

mental effects of galaxy formation as well as extend the dynamic range of distribution functions without the need to simulate large boxes. Another example is the use of FIRE-2 (Hopkins et al., 2018) physics model in Ma et al. (2018), to re-simulate various halos selected at $z = 5$ from dark matter only simulation boxes (largest box used is of side length 43 cMpc) at higher resolution. The re-simulated galaxies are combined with a weighting scheme based on the abundance of the target halos in the Universe, to produce composite distribution functions.

In this chapter we introduce FLARES, a suite of zoom resimulations during the EoR using the EAGLE model¹ to re-simulate a wide range of overdensities in the EoR. FLARES follows an approach similar to the GIMIC simulations to produce composite distribution functions. The EAGLE project (Schaye et al., 2015; Crain et al., 2015) is a suite of Smoothed Particle Hydrodynamics (SPH) simulations, calibrated to reproduce the stellar mass function and sizes of galaxies in the local Universe. EAGLE has been shown to be in good agreement with a large number of observables not used in the calibration (e.g. Lagos et al., 2015; Bahé et al., 2016; Furlong et al., 2017; Trayford et al., 2015, 2017; Crain et al., 2017). This includes predictions at high-redshift: Furlong et al. (2015) found reasonably good agreement with observationally inferred distribution functions of stellar mass and star formation rate out to $z = 7$. Unfortunately, there are very few well resolved galaxies in the fiducial EAGLE volume during the EoR. This is particularly the case for the most massive objects, which predominantly reside in protocluster environments, the progenitors of today’s collapsed clusters (Chiang et al., 2017; Lovell et al., 2018). FLARES allows us to significantly increase the number of galaxies simulated during the EoR with EAGLE. It also allows us to test the already incredibly successful EAGLE model in a new regime of extreme, high- z environments, whilst still resolving hydrodynamic processes at $10^6 M_\odot$ resolution, and provide predictions for a number of key upcoming observatories.

The purpose of this chapter is to introduce the resimulation method, our suite of zoom simulations, and present our predictions for the distribution of galaxies by stellar mass and star formation rate using the composite approach. We assume a Planck year 1 cosmology ($\Omega_0 = 0.307$, $\Omega_\Lambda = 0.693$, $h = 0.6777$, Planck Collaboration et al., 2014) and a Chabrier stellar initial mass function (IMF) throughout (Chabrier, 2003), and have corrected observational results accordingly.

¹ project website available at <https://flaresimulations.github.io/flares/>

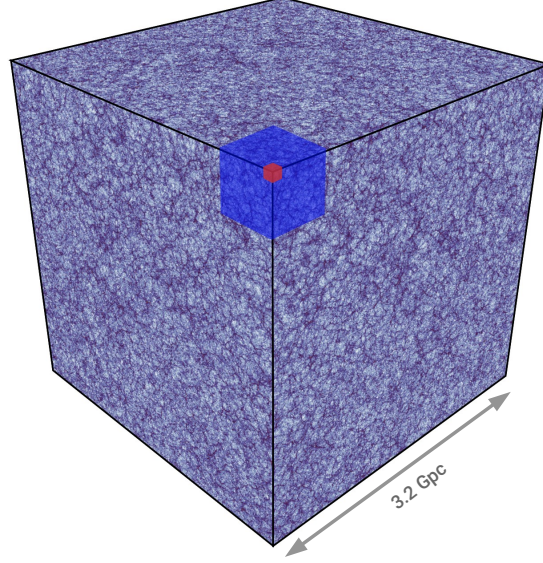


Figure 3.2: Diagram of the 3.2 cGpc box from which we select our regions (Barnes et al., 2017a). To demonstrate the increase in volume, we show the BLUE TIDES simulation ($L = 570$ cMpc; Feng et al., 2016) inset in blue, and the fiducial EAGLE simulation ($L = 100$ cMpc; Schaye et al., 2015) inset in red.

3.2 The FLARE Simulations

We will now detail our simulations, selection of the regions, the zoom resimulation technique, and our method for constructing composite distribution functions.

The EAGLE physics model has already been detailed in §1.2.3.1. Here we will touch upon the differences from the EAGLE reference volume in our adopted model. We use the AGNdT9 parameter configuration, which produces similar mass functions to the reference model but better reproduces the hot gas properties in groups and clusters (Barnes et al., 2017b). This is identical to that used in the C-EAGLE simulations, but differs from the fiducial Reference simulation (see §1.2.3.1). It uses a higher value for C_{visc} , which controls the sensitivity of the BH accretion rate to the angular momentum of the gas, and a higher gas temperature increase from AGN feedback, ΔT . These parameter changes impact the central black hole accretion, which has been shown to be efficient only at halo masses $> 10^{12} M_{\odot}$ (Bower et al., 2017). At $z = 10$ no FLARE galaxies reside in such halos, however at $z = 5$ a minority do ($< 0.2\%$), which may affect the early star formation histories of cluster galaxies (Bahé et al., 2017). The simulations have an identical resolution to the 100 cMpc Eagle Reference simulation box, with a dark matter and an initial gas particle mass of $m_{\text{dm}} = 9.7 \times 10^6 M_{\odot}$ and $m_{\text{g}} = 1.8 \times 10^6 M_{\odot}$ respectively, and has a gravitational softening length of 2.66 ckpc at $z \geq 2.8$.

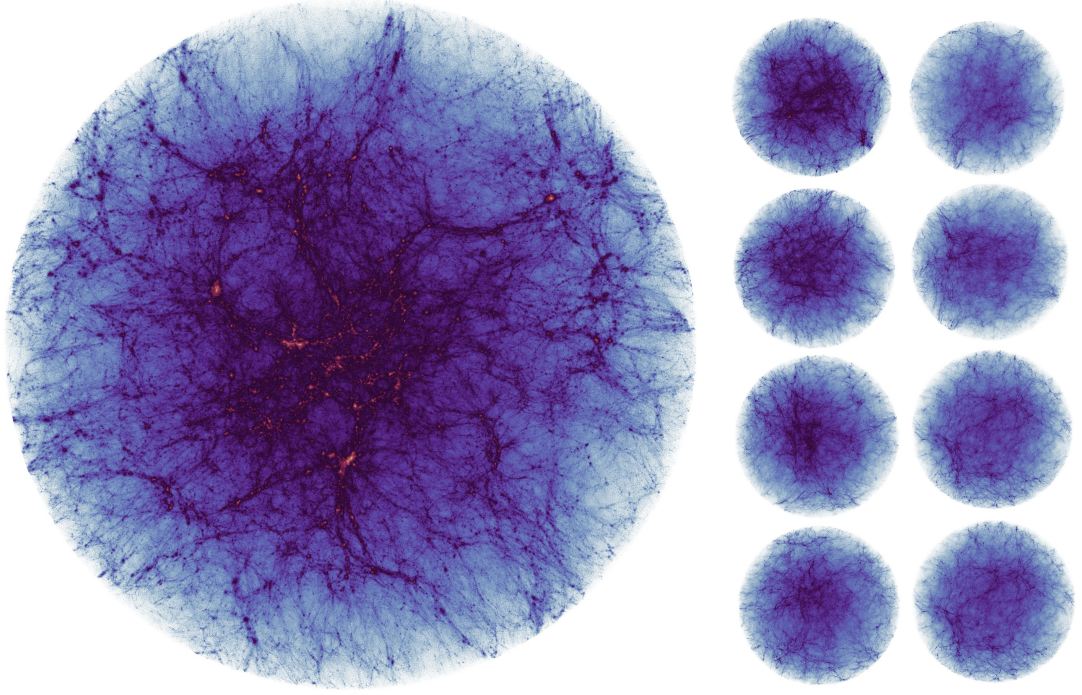


Figure 3.3: Visualisation of the dark matter integrated density in a number of resimulation regions of differing overdensity (δ), made with Py-SPHViewer (Benitez-Llambay, 2015). The region on the left shows the most overdense region (00, $\delta = 0.970$). The regions to the right are (anticlockwise from top left) 17, 20, 22, 24, 26, 28, 30, 38, with overdensities $\delta = [0.616, 0.266, 0.121, -0.007, -0.121, -0.222, -0.311, -0.479]$, respectively.

3.2.1 Region Selection

We use the same parent simulation as that used in the C-EAGLE simulations (Barnes et al., 2017a): a $(3.2 \text{ cGpc})^3$ dark matter-only simulation with a particle mass of $8.01 \times 10^{10} M_\odot$, using a Planck Collaboration et al. (2014) cosmology. Figure 3.2 shows a diagram of the box compared to the fiducial EAGLE reference volume, as well as the BLUETIDES simulation (Feng et al., 2016). The highest redshift snapshot available for this simulation is at $z = 4.67$, which we use for our selection. Within this snapshot, we select spherical volumes that sample a range of overdensities. By taking a sufficiently large radius we can ensure that the density fluctuations averaged on that scale are linear, such that the distortion in the shape of the Lagrangian volume during the simulation will not be too extreme and that the ordering of the density fluctuations is preserved. The regions, and their overdensities, are given in Table 3.1.

To determine the density, we first distribute the mass onto a high resolution, $3.2 \text{ cGpc} / 1200 \sim 2.67 \text{ cMpc}$ cubic grid using a nearest grid point assignment scheme. We then find the density on larger scales by convolving the grid with a spherical top-hat filter of radius $14 h^{-1} \text{ cMpc}$.² We

² Code provided at <https://github.com/christopherlovell/DensityGridder>

find, in test volumes, that this gives densities very close to those calculated from the raw particle data. The overdensity is then defined as

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x})}{\bar{\rho}} - 1, \quad (3.1)$$

where ρ is the density at grid coordinates \mathbf{x} , and $\bar{\rho}$ is the mean density in the box. The upper panel of Figure 3.4 shows the distribution of overdensity in log-space, alongside a fitted log-normal distribution.

We select regions for resimulation with two different goals: firstly, we select a number of regions of high overdensity in order to obtain a large sample of the first massive galaxies to form in the Universe; and secondly we select regions with a range of overdensities in order to explore the environmental impact (bias) on galaxy formation. In order to achieve the first goal we select the 16 most overdense regions in the volume, which have $\delta \geq 0.8$. For the second goal, we select two regions at each overdensity based on their *rms* overdensity σ , in the range $\sigma \in [4, 3, 2, 1, 0.5, 0, -0.5, -1, -2, -3]$. We choose two regions of each overdensity in order to minimise the effect of cosmic variance at fixed overdensity; we also select an additional two mean density regions, to increase the sampled volume of these common regions. Finally, we also select the two most underdense regions ($\delta \sim -0.45$) in order to cover the whole dynamic range. This gives a total of 40 regions. Figure 3.9 shows the GSMF for each region individually at $z = 5$. Whilst there is significant variation with overdensity ($\gtrsim 2$ dex near the knee), at fixed overdensity the scatter is low; this is particularly evident at high overdensities, where we selected a number of regions with very similar overdensity. This suggests that the effect of cosmic variance is low, and that the number of regions chosen was sufficient to demonstrate the trends presented in this article. However, we plan to run a greater number of simulations to further reduce the noise above the knee of the stellar mass function; an advantage of the resimulation approach is that this can simply be achieved by running more simulations to increase the total simulated volume.

The selected regions are listed in Appendix 3.A and the range of overdensities that each covers (evaluated at each point on the 2.67 cMpc grid enclosed by that volume) is shown in the lower panel of Figure 3.4. We discuss how to combine the resimulations so as to obtain a representative sample of the whole Universe in Section 3.2.3.

3.2.2 The Resimulation Method

Structures in FLARES, similar to the standard EAGLE analysis, are first found using a Friends-Of-Friends (FOF, Davis et al., 1985) finder, then split into bound substructures using the SUBFIND

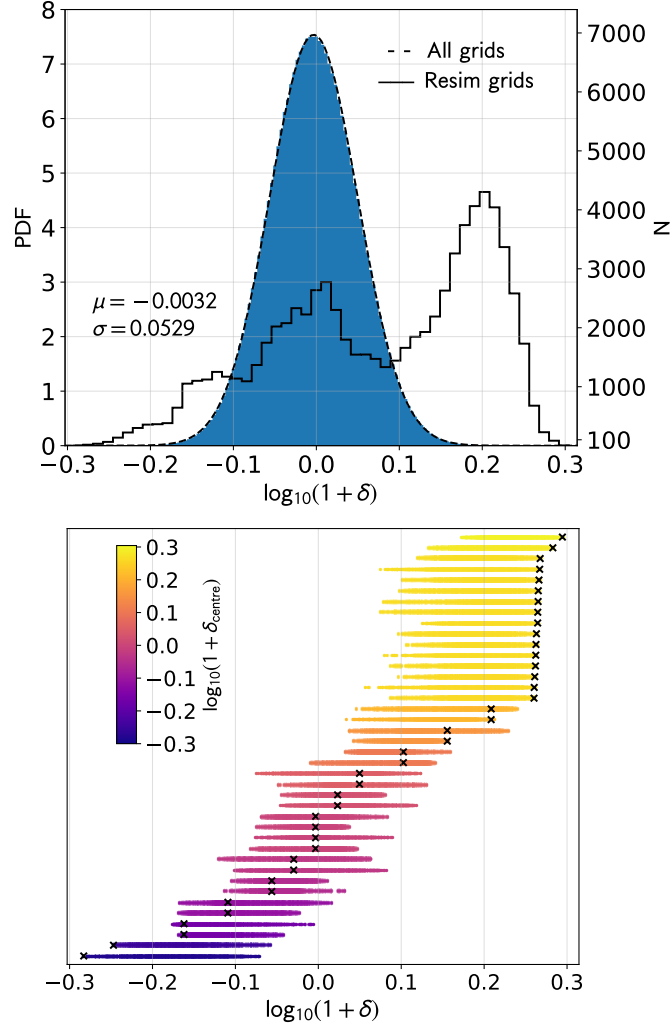


Figure 3.4: *Top panel*: the probability distribution function of sampled overdensities. The dashed black line shows a lognormal fit with the given parameters. The solid blue histogram shows the grid locations that lie within one of our resimulation volumes. The solid black histogram shows the distribution of our selected regions in overdensity, binned into 50 equal width bins, with the right y-axis showing only their number counts. *Bottom panel*: the distribution of overdensities within each simulation volume. The vertical displacement is arbitrary. The cross shows the overdensity measured at the centre of the resimulated volume and the spread of values shows the overdensities within each volume evaluated at each point on the 2.67 cMpc grid.

algorithm (Springel et al., 2001).³ Their properties are then defined using those stellar particles within 30 pkpc of the location of the most tightly-bound stellar particle. We limit our analysis to galaxies sampled by at least 50 star particles, which corresponds to a mass limit of approximately $\log_{10}(M_{\star} / M_{\odot}) \geq 7.95$.

Galaxies on the edge of the high resolution region will not be modelled correctly due to the

³ A number of galaxies identified by subfind are, on close inspection, ‘spurious’ structures, which manifest as an unrealistic ratio between the stellar, gas or dark matter components (see McAlpine et al., 2016, for a discussion). These galaxies make up less than 0.1% of all galaxies $> 10^8 M_{\odot}$ at $z = 5$, and are typically low mass. We use the following conditions to flag spurious galaxies: any subhalo with zero mass in the stellar, gas or dark matter components. Once these galaxies have been identified, we remove them from the subfind catalogues, and add their particle properties to the parent ‘central’ subhalo.

presence of a pressureless boundary. To avoid this, we resimulate a larger region, $1.25 \times 15h^{-1}$ cMpc in radius and ignore all galaxies with their centre of potential outside the inner $14h^{-1}$ cMpc. At higher redshift the Lagrangian high resolution region can deform, but we found that it is close to spherical out to the highest redshifts considered in this work ($z = 10$). In order to select only the inner $14h^{-1}$ cMpc in case of any deformation, we fit an equation of a sphere to the boundary dark matter particles identified using the module `ConvexHull` in the `PYTHON` package `scipy.spatial` to find the centre. Figure 3.3 shows the dark matter distribution within the cutout radius for a range of resimulations of differing overdensity, at $z = 4.7$. We also show the fiducial periodic EAGLE volume to provide a visual comparison of the differing environments probed.

3.2.3 Distribution Function Weighting

In this section we describe how we combine our resimulations to obtain a statistically-correct representation of the universal cosmological distribution of galaxies. As we show below in Section 3.3.2, distribution functions, such as the galaxy stellar mass function, vary with the overdensity of the resimulated volume. Therefore, it is necessary to *weight* each resimulation to reproduce the correct distribution of those overdensities averaged over the whole Universe, i.e. the cosmic mean.

As mentioned in Section 3.2.1, the overdensity within spherical top-hat regions of radius $14h^{-1}$ cMpc is sampled on a 2.67 cMpc grid; we label this sample δ_g . Since the grid sampling is finer than the size of the resimulation volume, each resimulation volume is associated with just under 2000 different values of δ_g . We show the distribution of those δ_g within each resimulation volume in the lower panel of Figure 3.4. The most overdense regions, whilst containing a single highly overdense point, in fact contain points covering a range of overdensities. It is, therefore, important to account for this spread in sampled overdensity, rather than just using the central overdensity when determining the contribution from any particular resimulation volume.

The top panel of Figure 3.4 contrasts the PDFs of δ_g for the whole box and for our resimulated sample. To generate the correct mean distributions, we divide into bins of overdensity as shown by the histogram in Figure 3.4 (black solid line), then *weight* the resimulations appropriately to reproduce the cosmic distribution. Specifically, we do the following:

- The overdensity domain is split up into 50 bins of equal width in $\log_{10}(1+\delta)$, $i = 1 \dots N_\delta$.⁴

For each of these, it is possible to assign a weight, $w_{\text{true},i}$, in proportion to the fraction of δ_g that lie in that bin, such that $\sum_i w_{\text{true},i} = 1$.

⁴ We tested using a greater number of bins and found that the quantitative weights did not change significantly.

- Each resimulation, j , is similarly distributed over these overdensity bins with weights, w_{ij} , in proportion to the enclosed values of δ_g . Thus $\sum_i w_{ij} = 1$.
- The sample weight associated with each bin is $w_{\text{sample},i} = \sum_j w_{ij}$.
- To obtain the correct universal average, we therefore have to weight each density bin by the ratio $r_i = w_{\text{true},i}/w_{\text{sample},i}$.

Ideally, we would associate each galaxy with the local value of δ_g . However, for the purposes of simplicity in this paper, we give all galaxies within a particular resimulation equal weight – this will give some dispersion over the more correct method, which we will implement in a future paper.

- Hence we adjust the contribution of each resimulation by a factor $f_j = \sum_i r_i w_{ij}$.

We note that

$$\begin{aligned} \sum_j f_j &= \sum_j \sum_i r_i w_{ij} = \sum_i r_i \sum_j w_{ij} \\ &= \sum_i r_i w_{\text{sample},i} = \sum_i w_{\text{true},i} = 1. \end{aligned} \quad (3.2)$$

These simulation weighting factors are listed in Table 3.1.

We further note that, at higher redshifts, the overdensities will evolve. Nevertheless, because even the most extreme perturbations are only mildly non-linear, we would expect that the ordering of the overdensities would largely be preserved. Hence, we use the same sampling at all redshifts. That also allows for a much more direct comparison of the evolution within each overdensity sample.

3.3 Results

3.3.1 Galaxy Number Counts

We begin by examining the raw number counts of galaxies. Figure 3.5 shows the cumulative distribution function of galaxies with stellar mass for both FLARES and the EAGLE Reference periodic volume ($V = (100 \text{ cMpc})^3$). We produce over ~ 20 times more $10^{10} M_\odot$ galaxies at $z = 5$ than obtained in the 100 cMpc periodic volume, despite the fact that the total high-resolution volume of all resimulated regions is only 50% larger than the periodic volume. This confirms that the first galaxies are significantly biased to higher overdensity regions.

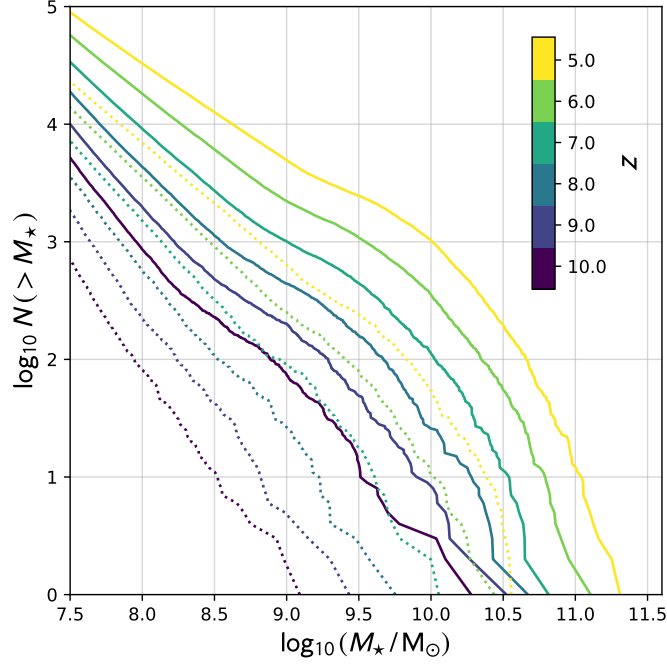


Figure 3.5: Cumulative distribution of stellar masses for all FLARES regions combined (solid) and the fiducial EAGLE Reference volume (dashed).

3.3.2 The Galaxy Stellar Mass Function

The Galaxy Stellar Mass Function (GSMF) describes the number of galaxies per unit volume per unit stellar mass interval $d\log_{10} M$,

$$\phi(M) = N / \text{Mpc}^{-3} \text{ dex}^{-1} , \quad (3.3)$$

and is commonly described using a Schechter function (Schechter, 1976),

$$\phi(M) d\log_{10} M = \ln(10) \phi^* e^{-M/M^*} \left(\frac{M}{M^*} \right)^{\alpha+1} , \quad (3.4)$$

which describes the high- and low-mass behaviour with an exponential and a power law dependence on stellar mass, respectively. Recent studies have found that a double Schechter function can better fit the full distribution (e.g. the GAMA survey, Baldry et al., 2008).

$$\phi(M) d\log_{10} M = \ln(10) e^{-M/M^*} \left[\phi_1^* \left(\frac{M}{M^*} \right)^{\alpha_1+1} + \phi_2^* \left(\frac{M}{M^*} \right)^{\alpha_2+1} \right] . \quad (3.5)$$

The low mass slope of the second Schechter function contributes to only a very narrow dynamic range. Above this range the exponential dominates, and below this the low mass slope of the first Schechter function dominates. It is therefore poorly constrained by the binned data, and so as not to introduce further degrees of freedom into our fit we fix it at $\alpha_2 = -1$. We define the

stellar mass M_\star as the total mass of all star particles, associated with the bound subhalo, within a 30 pkpc aperture centred on the potential minimum of the subhalo.⁵

3.3.2.1 The cosmic GSMF

In this section, we present results for the universal GSMF, averaged within our $(3.2 \text{ cGpc})^3$ box. This is obtained by combining the individual GSMFs from each of our resimulation volumes with appropriate weighting, as described in Section 3.2.3. This can be explained as follows and applicable to creating any other composite distribution function like the star formation rate function or luminosity functions,

$$\phi_i = \sum_j w_j N_{ij} / (V \Delta b), \quad (3.6)$$

where ϕ_i is the galaxy number density in bin ‘ i ’, w_j is the weight associated with the region ‘ j ’, N_{ij} is the number of galaxies associated with region ‘ j ’ in bin ‘ i ’, V is the volume of a single region and Δb is the bin width. Similarly the poisson error associated with a stellar mass bin, $\phi_{\text{err},i}$ can be expressed as

$$\phi_{\text{err},i} = \sqrt{\sum_j \left(w_j \sqrt{N_{ij}} \right)^2} / (V \Delta b). \quad (3.7)$$

The top panel of Figure 3.6 shows the GSMF for redshifts between $z = 10 \mapsto 5$. We show differential counts in bins 0.2 dex in width (with 1σ poisson uncertainties). The solid lines show double-Schechter function fits at each integer redshift. The normalisation increases with decreasing redshift, and the characteristic mass (or knee) of the distribution shifts to higher masses. This is more clearly seen in Figure 3.7, which shows the evolution of the double-Schechter parameters with redshift. The low-mass slope also gets shallower with decreasing redshift, from -3.5 at $z = 10$ to -2.0 at $z = 5$.

Our composite GSMF significantly extends the dynamic range of the GSMF compared to the periodic volumes. To demonstrate, the right panel of Figure 3.6 shows the FLARES double-Schechter fits, alongside the binned counts from the Reference periodic volume. At each redshift the maximum stellar mass probed is approximately an order of magnitude larger in FLARES. In fact, the periodic reference volume barely probes the exponential tail of the high mass component of the GSMF. When fitting a double-Schechter to the binned Reference volume counts we found that the parameters of the high mass component were completely unconstrained. However, it is

⁵ Two substructures within 30 pkpc of each other are still identified as separate structures, and only the particles associated with each structure contributes to its aperture-measured properties.

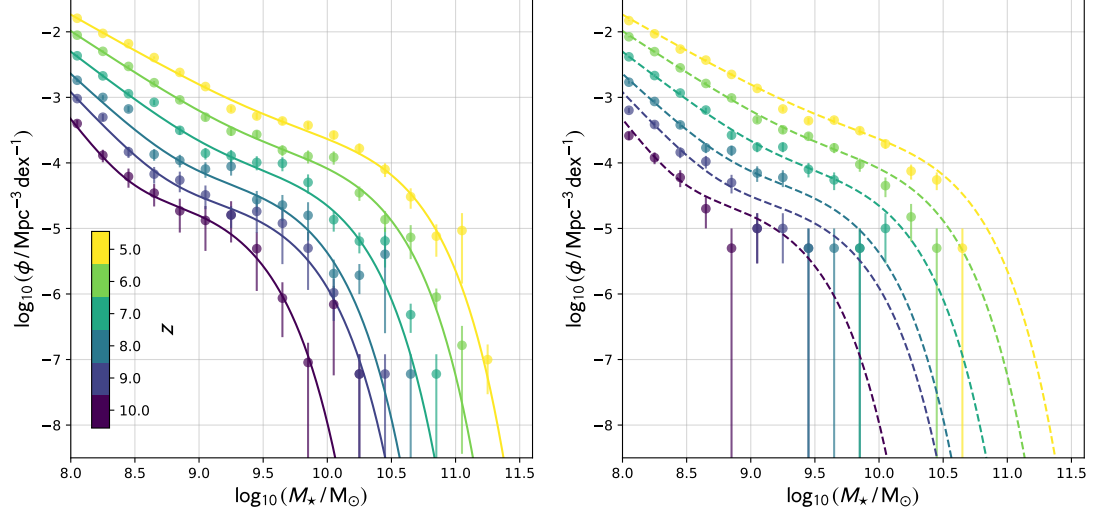


Figure 3.6: *Left:* Redshift evolution of the FLARES composite galaxy stellar mass function. Points show binned differential counts with Poisson 1σ uncertainties from the simulated number counts. Solid lines show double-Schechter function fits, quoted in Table 3.2. The parameter evolution is shown in Figure 3.7. *Right:* Same plot as the left panel, but points show the counts from the periodic Reference volume. The dashed lines show the double-Schechter fitted relation from FLARES. The coverage of the massive end in the periodic volume is poor.

clear from the bottom panel of Figure 3.6 that the low-mass slope is consistent between the Reference volume and FLARES. We have also tested that this is the case for the $(50 \text{ Mpc})^3$ AGNdT9 periodic volume. This provides evidence that our weighting method is accurately recovering the composite GSMF, without suffering from completeness bias. We note that the GSMF in the AGNdT9 and Reference periodic volumes is also in agreement at the low mass end, which gives us confidence that model incompleteness is not affecting our results.

In Figure 3.8 we show the composite FLARES GSMF against a number of high- z observational constraints in the literature (Gonzalez et al., 2011; Duncan et al., 2014; Song et al., 2016; Stefanon et al., 2017; Bhatawdekar et al., 2019). These studies show a spread of ~ 0.5 dex at $z = 5$, which highlights the difficulty of accurately measuring the GSMF at high redshift. The FLARES composite GSMF lies within this inter-study scatter, most closely following the relations derived by Song et al. (2016) up to $z = 7$. At $z \geq 8$ observational constraints are limited to cluster lensing studies such as the Hubble Frontier Fields, which do not probe the high-mass end due to the limited volume probed, but can reach very lower stellar masses ($\sim 10^7 M_\odot$). The fits presented in Bhatawdekar et al. (2019) have a higher normalisation than in FLARES over the accessible mass range, though they quote an uncertainty at $10^{8.5} M_\odot$ of ~ 0.6 dex at $z = 9$; FLARES lies within this uncertainty for the point sources, but is still in tension with the normalisation for disc-like sources.⁶ There is good agreement with the low-mass slope for both sources.

⁶ We show both disc-like and point-like constraints on the Bhatawdekar et al. (2019) GSMF; we note here that many of our galaxies have disc-like morphologies even at the highest redshifts.

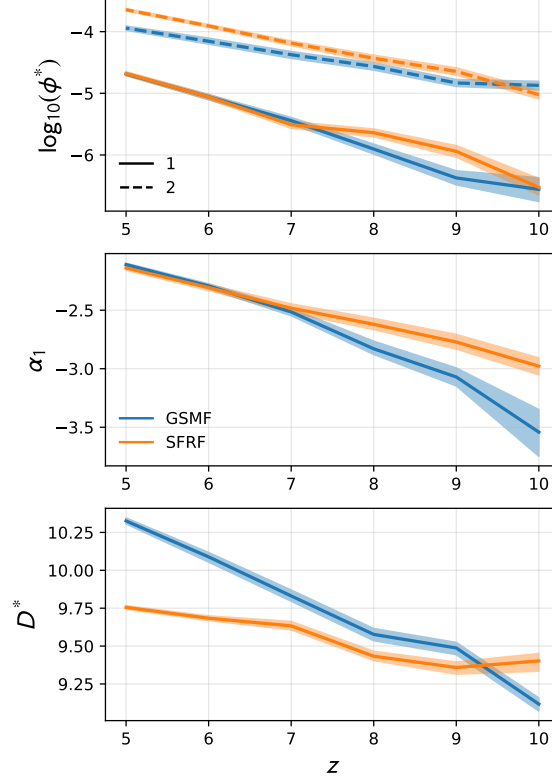


Figure 3.7: Parameter evolution for double-Schechter function fits to the FLARES composite galaxy stellar mass function (GSMF, blue) and star formation rate function (SFRF, orange). The low (1) and high (2) mass components are shown with solid and dashed lines, respectively. Shaded regions show the 16th – 84th percentile uncertainty obtained from the fit posteriors (see Appendix 3.B for details). The low-mass slope of the high-mass component (α_2) is fixed at -1. The characteristic mass of the GSMF (M_*) and the characteristic SFR of the SFRF (ψ_*) are shown in the bottom panel, labelled D_* . ψ_* is offset by $+10^8$ to aid comparison with M_* . The GSMF and SFRF show very similar behaviour; the normalisation of both components and the low-mass slope all increase with decreasing redshift. The characteristic mass increases with decreasing redshift for the GSMF, whereas the characteristic star formation rate of the SFRF shows a flatter redshift relation.

We also compare in Figure 3.8 to predictions from other galaxy formation models. The Feedback In Realistic Environments (FIRE) project performed zoom simulations of individual halos with masses between $10^8 - 10^{12} M_\odot$, which were then combined to provide a composite galaxy stellar mass function probing the low-mass regime (Ma et al., 2018). FLARES is consistent with FIRE at all redshifts where their mass range overlaps. Figure 3.8 also shows the GSMF from the 2015 and 2020 versions of L-GALAXIES (Henriques et al., 2015, 2020). Both models are in reasonably good agreement at all redshifts shown, but tend to underestimate the number density of massive galaxies at $z = 5$ compared to both FLARES and the observations.

Yung et al. (2019b) presented results from the Santa Cruz semi-analytic model (Somerville et al., 2015), which extends to a wide dynamic range. Whilst FLARES is consistent with this model for $z \leq 7$, at $z \geq 8$ the Santa Cruz model predicts a more power-law shape to the GSMF, with a

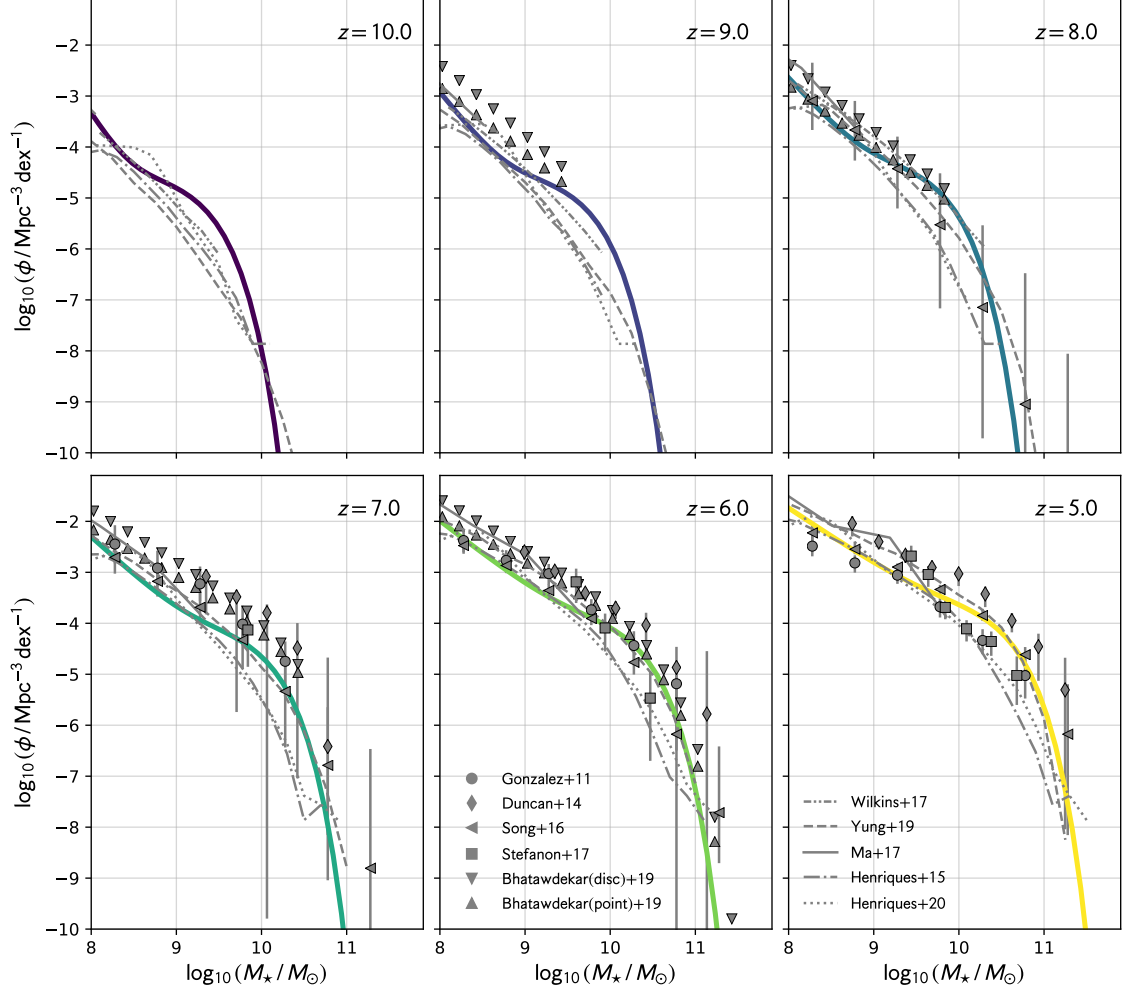


Figure 3.8: FLARES composite galaxy stellar mass function evolution, alongside observational constraints (Gonzalez et al., 2011; Duncan et al., 2014; Song et al., 2016; Stefanon et al., 2017; Bhatawdekar et al., 2019) as well as predictions from other models (Wilkins et al., 2017; Ma et al., 2018; Yung et al., 2019b; Henriques et al., 2015, 2020). There is some disagreement over the normalisation of the GSMF between different observational studies, however FLARES is consistent up to $z = 9$.

lower normalisation at the characteristic mass.

3.3.2.2 Environmental dependence of the GSMF

Our zoom simulations of a range of overdensities not only allow us to construct a composite GSMF for the entire $(3.2 \text{ Gpc})^3$ volume, but also investigate the environmental effect on the GSMF. Section 3.2 demonstrates the wide range of environments probed, from extremely underdense void regions, to the most overdense high redshift structures that are likely to collapse in to massive, $> 10^{15} M_{\odot}$ clusters by $z = 0$ (Chiang et al., 2013; Lovell et al., 2018).

Figure 3.9 shows the GSMF in bins of log-overdensity from $z = 5 - 9$. We use wider bins than previously (0.4 dex) due to the lower galaxy numbers in each resimulation. As expected, higher overdensity regions have a higher normalisation, $\sim +2$ dex above the lowest overdensity regions

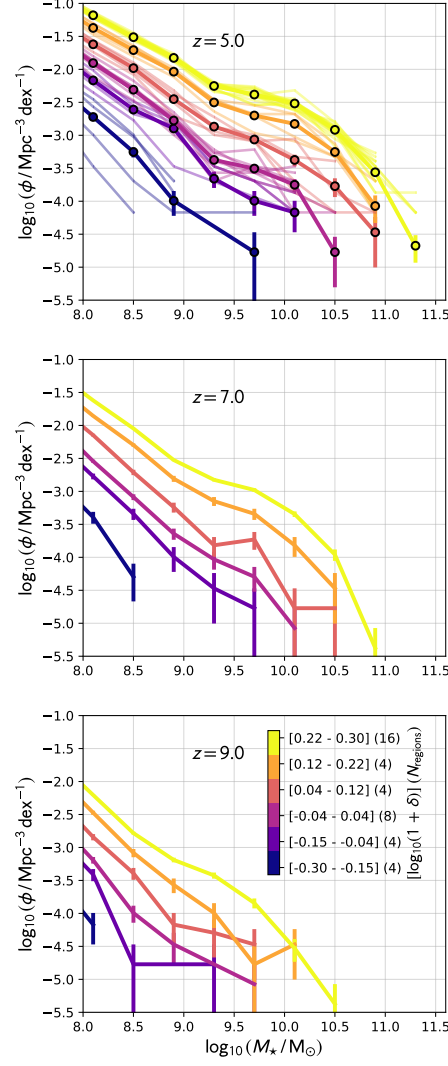


Figure 3.9: The FLARES GSMF for $z = 5, 7$ and 9 split by binned log-overdensity. The binning is shown in the legend, along with the number of regions in each bin. Poisson 1σ uncertainties are shown for each bin from the simulated number counts. The normalisation increases with increasing overdensity, and probes higher stellar masses. The top panel additionally shows the GSMF for each region individually (binned regions are highlighted using black edged markers).

at $M_\star/M_\odot = 10^{9.5}$ ($z = 5$). There is also an apparent difference in the shape as a function of log-overdensity: lower overdensity regions exhibit a distribution that is more power-law-like, whereas higher overdensity regions clearly show a double-Schechter-like knee. This may be due to the higher number of galaxies in the overdense regions, better sampling the knee, but may also point to differing assembly histories for galaxies in different environments. We will explore the star formation and assembly histories more closely in future work.

The dependence of the GSMF on overdensity may explain the tension between the composite FLARES GSMF and other models at $z > 7$ seen in Figure 3.8. Our much larger box allows us to sample extreme overdensities that are not present in smaller volumes. Double-Schechter forms of the GSMF at low- z have been attributed to the contribution of a passive and star forming

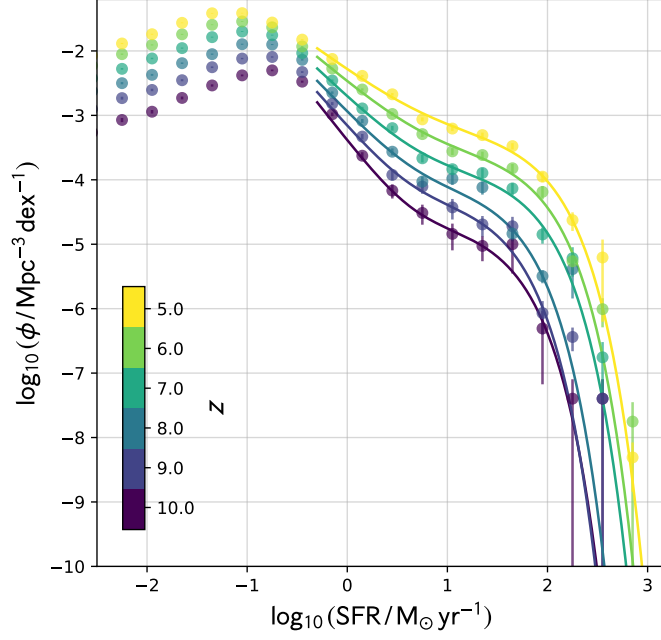


Figure 3.10: Redshift evolution of the FLARES composite star formation rate distribution function. Points show binned differential counts with Poisson 1σ uncertainties from the simulated number counts. Solid lines show double-Schechter function fits, quoted in Table 3.3.

population, each fit individually by a single Schechter function (Kelvin et al., 2014; Moffett et al., 2016), though this separation is not perfect (e.g. Ilbert et al., 2013; Tomczak et al., 2016). We suggest that the tension may be due to the small volume probed observationally at these depths, which does not probe extreme environments that contribute significantly to the cosmic GSMF.

3.3.3 The Star Formation Rate Distribution Function

The Star Formation Rate distribution Function (SFRF) describes the number of galaxies per unit volume per unit star formation rate interval $d\log_{10} \psi$, where ψ is the star formation rate,

$$\phi(\psi) = N / \text{Mpc}^{-3} \text{dex}^{-1} . \quad (3.8)$$

We define the SFR as the sum of the instantaneous SFR of all star forming gas particles, associated with the bound subhalo, within a 30 kpc aperture (proper) centred on the potential minimum of the subhalo.

3.3.3.1 The cosmic SFRF

In Figure 3.10 we plot the evolution of the FLARES composite SFRF. We provide counts in bins 0.3 dex in width. There is a clear low-mass turnover between $\sim 0.1 - 0.3 M_{\odot} \text{yr}^{-1}$, but above this the shape is well described by a double-Schechter function. Salim & Lee (2012) argue that a

single-Schechter is inadequate to describe the SFRF, as we find, though they propose a ‘Saunders’ function that does not provide a good fit to the FLARES SFRF. We provide fits using the following parametrisation,

$$\phi(\psi) d \log_{10} \psi = \ln(10) e^{-\psi/\psi^*} \times \left[\phi_1^* \left(\frac{\psi}{\psi^*} \right)^{\alpha_1+1} + \phi_2^* \left(\frac{\psi}{\psi^*} \right)^{\alpha_2+1} \right]. \quad (3.9)$$

We limit our fits to those galaxies with $\psi > 0.5 M_\odot \text{ yr}^{-1}$; these fits are provided in Table 3.3. We also plot the parameter evolution with redshift in Figure 3.7. The characteristic star formation rate, ψ_* , is offset by $+10^8$ to aid comparison with the GSMF characteristic mass, M_* .

The normalisation of both components (ϕ_1 ; ϕ_2), as well as the low-SFR slope (α_1), increase with decreasing redshift. These trends are surprisingly similar to those seen for the equivalent parameters in the GSMF. The low-SFR normalisation is almost identical, as is the high-SFR normalisation, with a small $\sim +0.2$ dex offset. The low-SFR slope α_1 is shallower than that of the GSMF at the highest redshifts ($z \geq 8$), but identical at lower redshifts. However, the evolution of the characteristic SFR is significantly flatter compared to that of the characteristic mass for the GSMF. This suggests a redshift-independent upper limit to the SFR.

This double-Schechter form of the SFRF is in some tension with observational constraints. Figure 3.11 shows a comparison with UV derived relations from Smit et al. (2012) and Katsianis et al. (2017) (the latter using Bouwens et al. 2015 data). For low-SFRs the observed normalisation is slightly higher (~ 0.3 dex) from $z = 5$ to 7. There is no prominent knee in the observed relations, and the exponential tail drops off at lower SFRs than in the simulations.

Figure 3.11 also shows results from recent cosmological models. As with the GSMF, there is some tension with the SFRF produced by the Santa Cruz models (Yung et al., 2019b). FLARES has a distinct double-Schechter shape, whereas the SC model appears as a single Schechter at $z = 5$, before evolving to a power law at $z = 10$. The BLUETIDES results (Wilkins et al., 2017) also show a similar power law relation at $z \geq 8$, in tension with the prominent knee in FLARES. Both L-GALAXIES models show similar power law-like behaviour, though with lower normalisation at the high-SFR end (Henriques et al., 2015, 2020), though in better agreement with the existing observational data at $z = 6$ compared to the Santa Cruz model and FLARES.

The offset in normalisation of the FLARES SFRF at high SFRs with the observations may be a selection effect due to highly dust-obscured galaxies. These galaxies, with number densities of $\sim 10^{-5} \text{ cMpc}^{-3}$ at $z \sim 2$ (Simpson et al., 2014), will be missed in higher redshift rest frame-UV observations. We will perform a direct comparison with the UV luminosity function, including

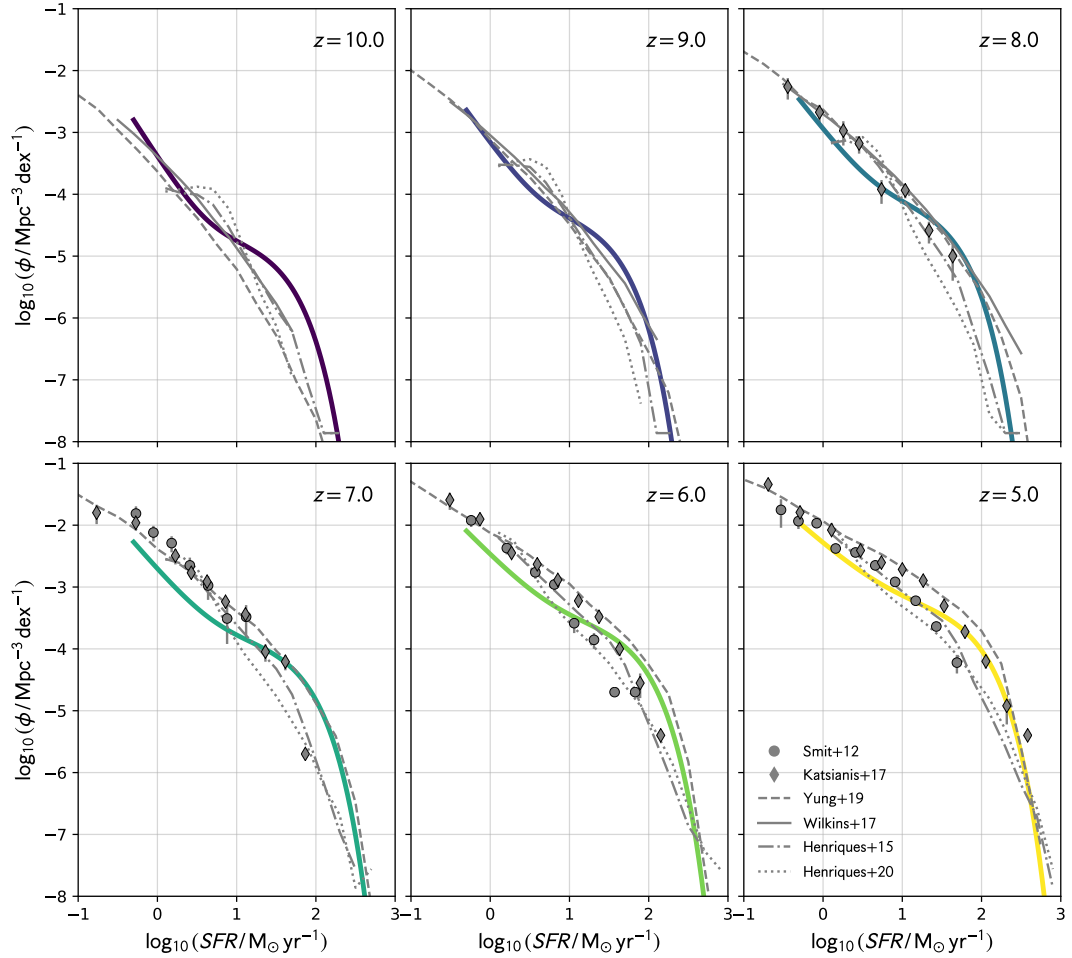


Figure 3.11: Evolution of the FLARES composite star formation rate distribution function (coloured, solid lines), compared with observational constraints from UV data and other model predictions. Smit et al. (2012) derive SFRs from UVLF data, as do Katsianis et al. (2017) using Bouwens et al. (2015) data. Both are corrected to a Chabrier IMF using the conversion factors quoted in Kennicutt Jr & Evans II (2012). The Santa-Cruz SAM (Yung et al., 2019b, dashed line) and BLUETIDES simulation (Wilkins et al., 2017) show a different behaviour, with a power law shape at higher redshifts, in contrast to the prominent knee seen in FLARES up to $z = 10$. Both L-GALAXIES models also show similar behaviour, though with lower normalisation at the high-SFR end (Henriques et al., 2015, 2020).

self-consistent modelling of dust attenuation, in Chapter 4 (which shows significant dust obscured SF to dominate at the high SFR end, see Figure 4.16). The offset may also be a modelling issue; EAGLE was not compared to high redshift observables during calibration, only to data at much lower redshifts ($z = 0.1$) than those studied here ($z \geq 5$). Improvements to the sub-grid modelling at high-redshift, particularly that of star-formation feedback, may improve the agreement.

To investigate what effect our sampling of highly overdense regions has on the composite shape of the SFRF, we now look at the overdensity dependence of the SFRF.

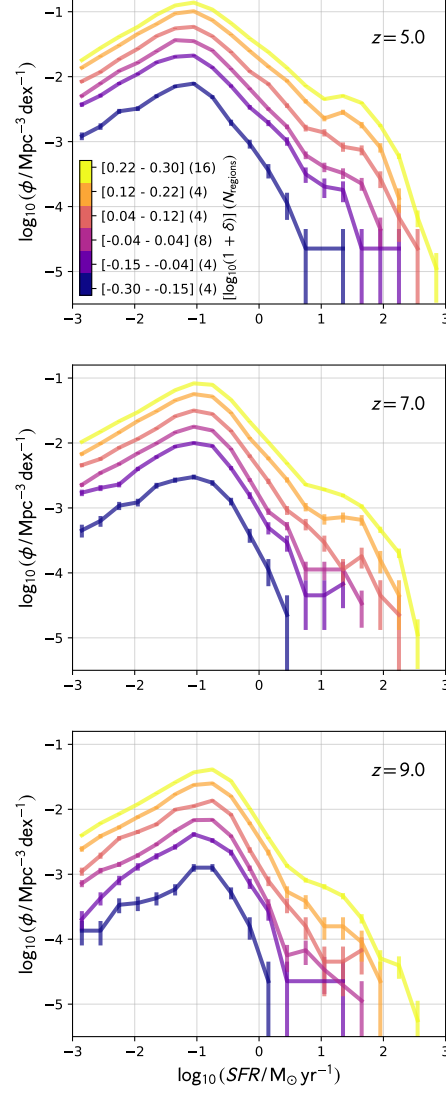


Figure 3.12: The FLARES SFRF for $z = 5, 7$ and 9 split by binned log-overdensity. The binning is shown in the legend, along with the number of regions in each bin. Poisson 1σ uncertainties are shown for each bin from the simulated number counts. The normalisation increases with increasing overdensity, and the maximum SFR increases.

3.3.3.2 Environmental dependence of the SFRF

Figure 3.12 shows the SFRF for regions binned by their log-overdensity. There is almost no variation in the shape as a function of overdensity except for the highest overdensities, which show a more prominent double-Schechter knee in the high-SFR regime. This behaviour is identical to that seen for the GSMF. This may explain why the shape of the FLARES composite SFRF differs with those of other cosmological models. FLARES better samples the rare, high-density regions that contribute significantly to the high-SFR ($\psi > 100 \text{ M}_{\odot} \text{ yr}^{-1}$) tail of the SFRF. Both BLUE TIDES and the Santa-Cruz model are run on regions with much smaller volumes (500^3 and 357^3 cMpc^3 , respectively), which may not probe the extreme regions sampled in the FLARES parent volume.

The mean density region in Figure 3.12 appears power law-like at all redshifts, which may present a better comparison with these models.

3.4 Conclusions

We have presented the first results from the FLARES simulations, resimulations with full hydrodynamics of a range of overdensities during the Epoch of Reionisation (EoR, $z \geq 5$) using the EAGLE (Schaye et al., 2015; Crain et al., 2015) physics. We described our novel weighting procedure that allows the construction of composite distribution functions that mimic extremely large periodic volumes, significantly extending the dynamic range without incurring prohibitively large computational expense. To demonstrate, we presented results for the galaxy stellar mass function (GSMF), the star formation rate distribution function (SFRF). Our findings are as follows:

- The FLARES GSMF exhibits a clear double-Schechter shape up to $z = 10$. Fits assuming this form show an increasing normalisation, shallower low-mass slope and higher characteristic turnover mass with decreasing redshift. The GSMF is in good agreement with observational constraints at all redshifts up to $z = 8$, at which point there is some tension at the knee of the distribution. The normalisation, and to a lesser extent the shape, of the GSMF shows a strong environmental dependence (i. e. bias).
- The SFRF also exhibits a clear double-Schechter shape in the high-SFR regime. As for the GSMF, the normalisation increases and the low-mass slope decreases with decreasing redshift; however the characteristic turnover mass varies only weakly with redshift. There is a mild tension with observational results, which tend to more closely resemble power law-like distributions. The SFRF shape and normalisation shows a similar environmental dependence to the GSMF.

Upcoming space based observatories, such as *JWST*, *Euclid* and *Roman Space Telescope* will provide further probes of the GSMF and SFRF up to $z = 10$. The large volumes probed by *Euclid* and *Roman Space Telescope* in particular will provide stronger constraints on those extreme galaxies that populate the high-mass / high-SFR tails of each distribution. Our weighting scheme provides a means of testing the latest, high resolution hydrodynamic simulations against such constraints. We will also be able to test the impact of cosmic variance on these large surveys.

3.A Selected regions

Table 3.1 lists the regions selected from the parent volume for resimulation.

3.B Fitted distribution functions

Table 3.2 and 3.3 show double-Schechter fit parameters to the GSMF and SFRF. We use `FitDF`, a python module for fitting arbitrary distribution functions using Markov Chain Monte Carlo (MCMC). `FitDF`⁷ is built around the popular `emcee` package (v3.0, Foreman-Mackey et al., 2013).

A Poisson form of the likelihood is typically used for distribution function analyses in Astronomy due to the relatively small number of observations. Due to our resimulation approach we cannot use this form of the likelihood, since the number counts obtained from the composite approach, scaled to the size of the parent box volume, significantly underestimate the errors. Instead, we use a Gaussian form for the likelihood,

$$\log(\mathcal{L}) = -\frac{1}{2} \left[\sum_i \frac{(N_{i,\text{obs}} - N_{i,\text{exp}})^2}{\sigma_i^2} + \log(\sigma_i^2) \right], \quad (3.10)$$

where the subscript i represents the bin of the property being measured, $N_{i,\text{obs}}$ is the inferred number of galaxies using the composite number density multiplied by the parent box volume, $N_{i,\text{exp}}$ is the expected number from the model, and σ_i is the error estimate. Using this form, σ can be explicitly provided from the resimulated number counts, $\sigma_i = N_{i,\text{obs}} / \sqrt{n_{i,\text{obs}}}$, where $n_{i,\text{obs}}$ is the number counts in bin i from the resimulations.

We use flat uniform priors in $\log_{10}(D^*)$, α_1 , $\log_{10}(\phi_1^*)$ and $\log_{10}(\phi_2^*)$. We fix $\alpha_2 = -1$ by setting a narrow top-hat prior around this value. We run chains of length 10^4 , then calculate the autocorrelation time, τ , on these chains (Goodman & Weare, 2010). We use τ to estimate the burn-in ($\tau \times 4$) and thinning ($\tau/2$) on our chains.⁸

⁷ The code can be found at <https://github.com/flaresimulations/fitDF>

⁸ The chains for each fit are available at <https://flaresimulations.github.io/flares/data.html>.

Table 3.1: Regions selected from the parent volume for resimulation. We provide their positions within the parent volume, their overdensity δ as defined by Equation 3.1, their *rms* overdensity σ , and weights, f_j , calculated as per Section 3.2.3.

index	(x, y, z)/(h ⁻¹ cMpc)	δ	σ	f_j
0	(623.5, 1142.2, 1525.3)	0.970	5.62	0.000027
1	(524.1, 1203.6, 1138.5)	0.918	5.41	0.000196
2	(54.2, 1709.6, 571.1)	0.852	5.12	0.000429
3	(153.6, 1762.0, 531.3)	0.849	5.11	0.000953
4	(39.8, 1686.1, 1850.6)	0.846	5.09	0.000444
5	(847.6, 1444.0, 1062.6)	0.842	5.07	0.000828
6	(1198.2, 135.5, 1375.3)	0.841	5.07	0.000666
7	(1012.0, 1514.4, 1454.8)	0.839	5.06	0.001178
8	(591.0, 359.6, 1610.2)	0.839	5.06	0.000265
9	(746.4, 820.5, 945.2)	0.833	5.03	0.001029
10	(1181.9, 1171.1, 974.1)	0.830	5.02	0.000387
11	(38.0, 670.5, 47.0)	0.829	5.02	0.000719
12	(1989.7, 368.7, 2076.5)	0.828	5.01	0.000668
13	(1659.0, 1306.6, 760.8)	0.824	4.99	0.000488
14	(57.8, 883.7, 2098.2)	0.821	4.98	0.001190
15	(609.0, 2018.6, 115.7)	0.820	4.98	0.000757
16	(122.9, 1124.1, 1304.8)	0.616	4.00	0.003738
17	(1395.2, 415.7, 1575.9)	0.616	4.00	0.004678
18	(128.3, 216.9, 258.4)	0.431	3.00	0.009359
19	(1400.6, 1686.1, 806.0)	0.431	3.00	0.012324
20	(699.4, 1760.2, 1725.9)	0.266	2.00	0.029311
21	(1951.8, 2022.3, 1709.6)	0.266	2.00	0.027954
22	(755.4, 1122.3, 867.5)	0.121	1.00	0.057876
23	(516.9, 325.3, 603.6)	0.121	1.00	0.062009
24	(937.9, 1382.5, 1077.1)	-0.007	0.00	0.074502
25	(1675.3, 1492.8, 1335.5)	-0.007	0.00	0.080377
26	(1270.5, 518.7, 862.0)	-0.121	-1.00	0.063528
27	(242.2, 1881.3, 1624.7)	-0.121	-1.00	0.058231
28	(1454.8, 1720.5, 1608.4)	-0.222	-2.00	0.034467
29	(430.1, 296.4, 359.6)	-0.222	-2.00	0.024216
30	(1733.1, 1097.0, 1060.8)	-0.311	-3.00	0.012087
31	(1821.7, 947.0, 1431.3)	-0.311	-3.00	0.013127
32	(1913.8, 1033.7, 45.2)	-0.066	-0.50	0.064280
33	(2009.6, 2024.1, 1693.4)	-0.066	-0.50	0.066277
34	(339.8, 934.3, 1646.4)	-0.007	0.00	0.076001
35	(1693.4, 914.5, 1977.1)	-0.007	-0.00	0.076486
36	(778.9, 900.0, 1866.8)	0.055	0.50	0.070408
37	(1790.9, 1239.7, 1765.6)	0.055	0.50	0.062451
38	(2078.3, 77.7, 141.0)	-0.479	-5.29	0.002721
39	(818.7, 110.2, 1628.3)	-0.434	-4.61	0.003366

z	M*	$\log_{10}(\phi_1^*/(\text{Mpc}^{-3} \text{ dex}^{-1}))$	$\log_{10}(\phi_2^*/(\text{Mpc}^{-3} \text{ dex}^{-1}))$	α_1
10	$9.117^{+0.041}_{-0.045}$	$-6.557^{+0.188}_{-0.197}$	$-4.871^{+0.065}_{-0.07}$	$-3.542^{+0.193}_{-0.206}$
9	$9.488^{+0.036}_{-0.044}$	$-6.372^{+0.116}_{-0.112}$	$-4.832^{+0.056}_{-0.057}$	$-3.07^{+0.076}_{-0.077}$
8	$9.577^{+0.039}_{-0.041}$	$-5.904^{+0.081}_{-0.08}$	$-4.565^{+0.059}_{-0.058}$	$-2.83^{+0.065}_{-0.048}$
7	$9.831^{+0.039}_{-0.035}$	$-5.443^{+0.051}_{-0.054}$	$-4.374^{+0.052}_{-0.059}$	$-2.515^{+0.03}_{-0.032}$
6	$10.089^{+0.029}_{-0.035}$	$-5.057^{+0.036}_{-0.047}$	$-4.156^{+0.05}_{-0.046}$	$-2.293^{+0.019}_{-0.023}$
5	$10.326^{+0.019}_{-0.02}$	$-4.686^{+0.023}_{-0.024}$	$-3.942^{+0.033}_{-0.034}$	$-2.11^{+0.012}_{-0.011}$

Table 3.2: Best fitting double-Schechter function parameter values for the Galaxy Stellar Mass Function. α_2 is fixed at -1 .

z	SFR*	$\log_{10}(\phi_1^*/(\text{Mpc}^{-3} \text{ dex}^{-1}))$	$\log_{10}(\phi_2^*/(\text{Mpc}^{-3} \text{ dex}^{-1}))$	α_1
5	$1.402^{+0.049}_{-0.067}$	$-6.525^{+0.142}_{-0.123}$	$-5.022^{+0.07}_{-0.069}$	$-2.978^{+0.071}_{-0.074}$
5	$1.359^{+0.036}_{-0.044}$	$-5.941^{+0.093}_{-0.093}$	$-4.645^{+0.058}_{-0.058}$	$-2.772^{+0.064}_{-0.06}$
5	$1.433^{+0.032}_{-0.028}$	$-5.639^{+0.059}_{-0.066}$	$-4.431^{+0.049}_{-0.058}$	$-2.62^{+0.051}_{-0.045}$
5	$1.633^{+0.03}_{-0.027}$	$-5.509^{+0.052}_{-0.057}$	$-4.186^{+0.036}_{-0.04}$	$-2.482^{+0.036}_{-0.038}$
5	$1.684^{+0.015}_{-0.015}$	$-5.059^{+0.041}_{-0.039}$	$-3.907^{+0.024}_{-0.026}$	$-2.307^{+0.026}_{-0.025}$
5	$1.755^{+0.011}_{-0.012}$	$-4.68^{+0.033}_{-0.033}$	$-3.644^{+0.02}_{-0.02}$	$-2.139^{+0.02}_{-0.019}$

Table 3.3: Best fitting double-Schechter function parameter values for the Star Formation Rate distribution function. α_2 is fixed at -1 .

FLARES II: THE PHOTOMETRIC PROPERTIES OF HIGH-REDSHIFT GALAXIES

4

In this Chapter, we present the photometric properties of galaxies in the First Light and Reionisation Epoch Simulations (FLARES). FLARES predicts a significantly larger number of intrinsically bright galaxies, which can be explained through a simple model linking dust-attenuation to the metal content of the interstellar medium, using a line-of-sight (LOS) extinction model. With this model we present the photometric properties of the FLARES galaxies for $z \in [5, 10]$. We show that the ultraviolet (UV) luminosity function (LF) matches the observations at all redshifts. The function is fit by Schechter and double power-law forms, with the latter being favoured at these redshifts by the FLARES composite UV LF. We also present predictions for the UV continuum slope as well as the attenuation in the UV. The impact of environment on the UV LF is also explored, with the brightest galaxies forming in the densest environments. We then present the line luminosity and equivalent widths of some prominent nebular emission lines arising from the galaxies, finding rough agreement with available observations. We also look at the relative contribution of obscured and unobscured star formation, finding comparable contributions at these redshifts.

4.1 The FLARE Simulations

The FLARE simulation strategy has already been explained in Chapter 3. Here we will detail the selection of our galaxy sample, their physical properties and our spectral energy distribution modelling technique.

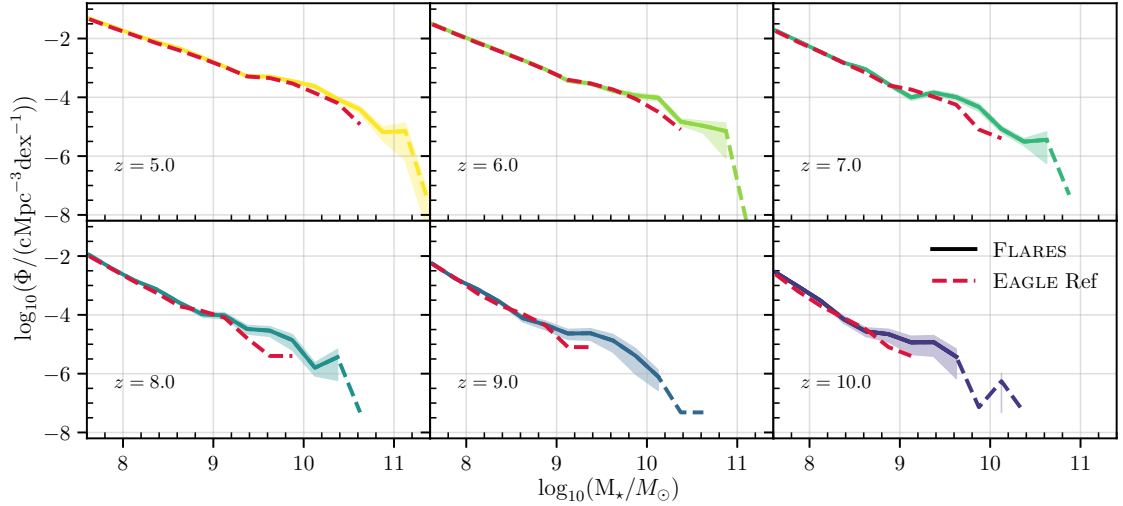


Figure 4.1: FLARES composite galaxy stellar mass function (black solid, dashed for bins with less than 5 galaxies) for $z \in [5, 10]$. Shaded regions denote the Poisson 1σ uncertainties for each bin from the simulated number counts for the FLARES galaxies. For comparison the GSMF from the 100 cMpc EAGLE Reference simulation box is shown in red.

4.1.1 Galaxy Identification

Galaxies in FLARES, similar to the standard EAGLE are identified with the SUBFIND algorithm (Springel et al., 2001; Dolag et al., 2009), which runs on bound groups found from via the Friends-Of-Friends algorithm (FOF, Davis et al., 1985). The stellar masses are defined using star particles within a 30 pkpc aperture centred on the most bound particle of the self-bound substructures. In this work, we concentrate on a broader definition of a galaxy with respect to Lovell et al. (2021a) (or Chapter 3), where only galaxies with a stellar mass $\gtrsim 10^8 M_\odot$ were considered in the analysis. Here we focus on objects with a combined total of more than 100 gas and star particles. This extends the stellar mass function down to $\sim 10^{7.5} M_\odot$ at $z = 5$.

FLARES has more than ~ 20 times the number of galaxies with a mass greater than $10^{10} M_\odot$ at $z = 5$ compared to the EAGLE reference volume (Schaye et al., 2015, see Figure 3.5). In Figure 4.1, we compare the galaxy stellar mass function of the galaxies in FLARES and the 100 cMpc EAGLE Reference simulation box. It can be seen that FLARES extends the range by at least an order of magnitude at the high-mass end compared to EAGLE.

4.1.2 Metal Content

Stellar evolution enriches galaxies with metals. This is governed by the rate at which stars are formed and the various mass loss events associated with their evolution (e. g. stellar winds, supernova explosion). The next generation of stars form from this enriched gas and evolve, con-

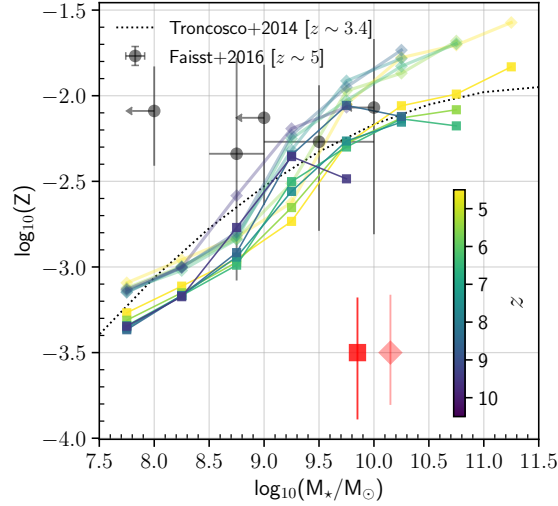


Figure 4.2: Mass weighted metallicities of the gas (darker square points) and stars (lighter diamond points) of the FLARE galaxies at $z \in [5, 10]$. Only the weighted median of the bins containing more than 5 galaxies are shown, with the maximum of the 16th and 84th percentile spread in the bins of the two data shown in red. The observational constraints on the gas-phase metallicity from Troncoso et al. (2014) at $z \sim 3.4$ and Faisst et al. (2016) at $z \sim 5$ are shown. Observational measurements of the stellar mass assume a Chabrier (2003) initial mass function with metallicities converted to a mass-fraction assuming $12 + \log_{10}(\text{O}/\text{H})_{\odot} = 8.69$ and $Z_{\odot} = 0.02$.

tinuing the cycle of metal enrichment in the galaxy. We show this process in Figure 4.2, where the evolution of the mass-weighted stellar and gas-phase metallicities are plotted as a function of galaxy stellar mass. The metallicity of galaxies generally increases with stellar mass. There is little evolution in the metallicity across redshifts, but a strong evolution with stellar mass by approximately an order of magnitude increase from the lowest to the highest stellar mass bin. The normalisation, as well as the trend in the metallicity with stellar mass, is similar to observed gas-phase metallicity seen in Troncoso et al. (2014) at $z \sim 3.4$, obtained using optical strong line diagnostics with the R_{23} parameter (for a summary see Kewley & Ellison, 2008). A similar normalisation of the relation at higher metallicities is seen at $z \sim 5$ in Faisst et al. (2016) using strong optical emission lines. It should be noted that the uncertainties on the observed metallicities is very large, due to the difficulty in measuring the value at $z \geq 5$. Observations from the upcoming *JWST* will be able to put tighter constraints in the high-redshift regime.

4.1.3 Spectral Energy Distribution Modelling

In this section, we detail the spectral energy distribution (SED) modelling of each galaxy. In this work, we model only the emission from stars (including reprocessing by gas and dust) and defer the treatment of accretion on to super-massive black holes to a future work. We broadly follow the approach implemented by Wilkins et al. (2016a, 2017, 2018, 2020) albeit with modifications to the dust modelling as described in §4.1.4.

4.1.3.1 Stellar Emission

We begin by modelling the pure stellar emission produced by each galaxy. To do this we associate each star particle with a stellar SED according to its age and metallicity (i. e. a simple stellar population or SSP). Throughout this work we utilise v2.2.1 of the Binary Population and Spectral Synthesis (BPASS) stellar population synthesis (SPS) models (Stanway & Eldridge, 2018) and assume a Chabrier initial mass function (IMF) throughout (Chabrier, 2003). As explored in Wilkins et al. (2016a, 2017, 2018, 2020) the choice of SPS and IMF can have a large effect on resulting broadband luminosities and emission line quantities.

4.1.3.2 Nebular Emission

Young stellar populations produce significant Lyman-continuum (LyC) emission. To account for the reprocessing of these photons by surrounding gas we associate each young ($t < 10$ Myr) star particle with a surrounding HII region (or birth cloud) powered by its LyC emission. To calculate the nebular emission we follow the approach detailed in Wilkins et al. (2020). In short, the pure stellar spectrum of each star particle is input to the CLOUDY (Ferland et al., 2017) photo-ionisation code. The metallicity of the associated HII region is assumed to be identical to the star particle, and we adopt the same dust depletion factors and relative abundances as Gutkin et al. (2016). We assume a reference ionisation parameter (defined at $t = 1$ Myr and $Z = 0.02$) of $\log_{10} U_{S,\text{ref}} = -2$, a hydrogen density of $\log_{10}(n_H/\text{cm}^{-3}) = 2.5$, and adopt CLOUDY's default implementation of Orion-type graphite and silicate grains.

4.1.4 Dust Attenuation

One of the most important ingredients in generating mock observations involves modelling the attenuation by dust. It has a major impact on the observed properties of galaxies, with almost 30% of all photons in the Universe having been reprocessed by dust grains at some point in their lifetime (Bernstein et al., 2002). There have been a few studies that have incorporated dust creation and destruction self-consistently into hydrodynamical simulations (e. g. Aoyama et al., 2017; McKinnon et al., 2017; Gjergo et al., 2018; Li et al., 2019; Graziani et al., 2020). They have found mixed success in matching many of the observed galaxy properties like the dust-to-stellar mass ratio, the dust-to-gas ratio or the dust-to-metal ratio. Many of these simulations also have information on the grain sizes or the contribution of different dust species to the total dust mass. This additional information can eliminate some of the post-processing assumptions involved in deriving observed properties (e. g. Hou et al., 2017; McKinnon et al., 2018; Kannan et al., 2019;

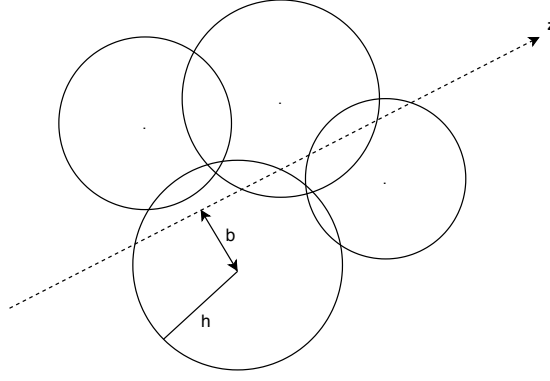


Figure 4.3: Line of sight tracing of the SPH density field, with the circles representative of SPH particles. h and b denote the smoothing length of the corresponding gas particle and the impact parameter to the LOS ray respectively.

Hirashita & Murga, 2020). However they also involve additional subgrid recipes which are poorly understood, and can get computationally intensive depending on the modelling techniques. A simple alternative is to model the effect of dust based on the properties of the existing stars and gas particles in the simulation. This is usually done by using the metallicity information of the ISM to build a model to attenuate the stellar spectra. They still incorporate information on the spatial distribution of dust and are therefore more detailed than a simple screen model.

In this work, for estimating the dust attenuation, each star particle is treated as a point in space with its emitted light reaching the observer through the intervening gas particles. We fix the viewing angle to be along the z -axis. For the purpose of this study we link the metal column density ($\Sigma(x, y)$), integrated along the LOS (z -axis in this case) to the dust optical depth in the V-band (550nm) due to the intervening ISM, $\tau_{\text{ISM,V}}(x, y)$, with a similar approach as in Wilkins et al. (2017). This relation can be expressed as

$$\tau_{\text{ISM,V}}(x, y) = \text{DTM} \kappa_{\text{ISM}} \Sigma(x, y), \quad (4.1)$$

where DTM is the dust-to-metal ratio of the galaxy and κ_{ISM} is a normalisation parameter which we have chosen to match the rest-frame far-UV (1500Å) luminosity function to the observed UV luminosity function from Bouwens et al. (2015) at $z = 5$. The DTM value of a given galaxy comes from the fitting function presented in Vijayan et al. (2019) (or Chapter 2, Equation 2.15), which is a function of the mass-weighted stellar age and the gas-phase metallicity. This allows for a varying DTM ratio across different galaxies as well as evolution across redshift as seen in observational works (e.g. De Vis, P. et al., 2019), depending on their evolutionary stage. This provides a single DTM value per galaxy, assuming no spatial variation. κ_{ISM} acts as a proxy for the properties of dust, such as the average grain size, shape, and composition. In a companion work, we will explore the impact of a range of different modelling approaches.

$\Sigma(x, y)$ is obtained by integrating the density field of particles along the z-axis with the smoothing kernel of the SPH particle. FLARES uses the same flavour of SPH used by EAGLE, ANARCHY (see Schaller et al., 2015, for more details). The kernel function can be expressed as follows:

$$W(r, h) = \frac{21}{2\pi h^3} \begin{cases} (1 - \frac{r}{h})^4 (1 + 4\frac{r}{h}) & \text{if } 0 \leq r \leq h \\ 0 & \text{if } r > h, \end{cases} \quad (4.2)$$

where h is the smoothing length of the corresponding particle and r is the distance from the centre of the particle. The smoothed density line integral across a particular particle can be calculated by using the impact parameter, b which is calculated from the centre of the particle (illustrated in Figure 4.3). Using the impact parameter of every gas particle in front of the selected stellar particle, the LOS metal column density can be calculated as follows:

$$\Sigma(x, y) = 2 \sum_i Z_i m_i \int_0^{\sqrt{h_i^2 - b_i^2}} W(r, h_i) dz; \quad r^2 = b_i^2 + z^2, \quad (4.3)$$

where the index i denotes gas particles along the LOS, with Z and m the metallicity and mass of the particle respectively. To simplify this calculation, impact parameters can be normalised with the smoothing length, and thus generate pre-computed values of the LOS metal density which can be readily used to compute the density for arbitrary values of smoothing length and impact parameters.

Other than the dust extinction along the LOS, there is an additional component of dust that affects young stellar populations that are still embedded in their birth cloud. Effect of the birth cloud attenuation in our galaxies is a phenomenon that happens below the resolution scale, since stellar clusters form on sub-kpc scales. The birth cloud dust optical depth in the V-band for our model can be expressed in a similar manner to equation 4.1 as

$$\tau_{\text{BC,V}}(x, y) = \begin{cases} \kappa_{\text{BC}}(Z/0.01) & t \leq 10^7 \text{yr} \\ 0 & t > 10^7 \text{yr}, \end{cases} \quad (4.4)$$

where κ_{BC} just like κ_{ISM} , is a normalisation factor, which also encapsulate the dust-to-metal ratio in the stellar birth clouds. This implies that we assume a constant dust-to-metal ratio in birth clouds for all galaxies. Here, Z is the metallicity of the stellar particle with age less than 10^7 yr, following the assumption from Charlot & Fall (2000) that birth clouds disperse on these timescales. Hence, only young stellar particles are affected by this additional attenuation. With these parameters the optical depth in the V-band is linked to other wavelengths using a simple simple power-law relation

$$\tau_\lambda = (\tau_{\text{ISM}} + \tau_{\text{BC}}) \times (\lambda/550\text{nm})^{-1}. \quad (4.5)$$

This functional form yields an extinction curve flatter in the UV than the Small Magellanic Cloud curve (Pei, 1992), but not as flat as the Calzetti et al. (2000) curve.

As discussed earlier there are two free parameters in our model, κ_{ISM} that links the optical depth in the ISM to the LOS metal surface density and κ_{BC} linking the stellar particle metallicity to the optical depth due to the presence of a birth cloud in young stellar populations. To obtain the values for these parameters we do a simple grid search approach. We make an array of candidate κ_{BC} values in the closed range $[0.001, 2.]$. For each κ_{BC} , we generate the UV LF for a grid of κ_{ISM} values in the range $(0, 1]$ at $z = 5$. These are then compared to the Bouwens et al. (2015) UV LF at $z = 5$ using a simple chi-square analysis to obtain the corresponding value for κ_{ISM} (only $M_{\text{UV}} < -18$ is used). We then generate the corresponding UV-continuum slope (β) as well as the $[\text{OIII}]\lambda_{4959,5007}$ and $\text{H}\beta$ line luminosity and equivalent widths (EW) for a given combination of $(\kappa_{\text{BC}}, \kappa_{\text{ISM}})$. The combination of $(\kappa_{\text{BC}}, \kappa_{\text{ISM}})$ value that best matches the $M_{\text{UV}} - \beta$ observations from Bouwens et al. (2012, 2014) at $z = 5$ (Figure 4.18) and the $[\text{OIII}]\lambda_{4959,5007} + \text{H}\beta$ line luminosity and EW relations versus UV luminosity and stellar mass at $z = 8$ from De Barros et al. (2019) (Figure 4.19) is chosen as our default model. This process leads a value of $\kappa_{\text{BC}} = 1$ and $\kappa_{\text{ISM}} = 0.0795$, which is used for all redshifts considered in this study. A higher value for κ_{BC} is favoured to get better agreement with the β observations while the line luminosity and EW relations prefer a lower value. Hence the chosen value of κ_{BC} is a way to incorporate the effect of both these observations. Future measurements in this observational space from current and upcoming telescopes, would help to further tighten our constraints on this value. The parameter search is explained further in Appendix 4.A. It is worth noting that by using fixed choice of these parameters, we assume there is no evolution in the general properties of the dust grains in galaxies such as the average grain size, shape, and composition.

We also show in Appendix 4.C how some of the observables presented in the next sections change on using different extinction curves available from literature.

4.2 Photometric Properties

4.2.1 UV Luminosity Function

The UV LF evolution of high-redshift galaxies is a parameter space where there are numerous observational studies (e. g. Bunker et al., 2004; Bouwens et al., 2006; Wilkins et al., 2011a; Bouwens et al., 2015; Finkelstein et al., 2015). We begin by calculating the rest-frame UV LF of the FLARES galaxies.

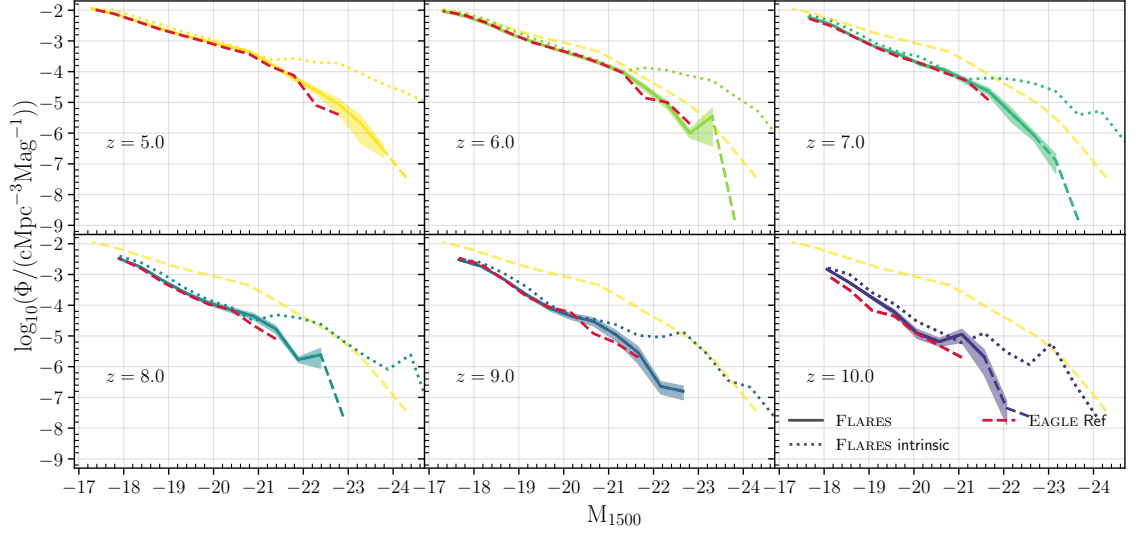


Figure 4.4: FLARES composite intrinsic (dotted) and dust attenuated (solid, dashed for bins with less than 5 galaxies) UV LF for galaxies in $z \in [5, 10]$. Shaded region denote the Poisson 1σ uncertainties for each bin from the simulated number counts for the dust attenuated UV LF. For comparison the dust attenuated UV LF from the EAGLE Reference volume is plotted in red. We also plot the $z = 5$ dust attenuated UV LF (dashed line) alongside other redshifts to aid comparison.

4.2.1.1 LF creation

Unlike cosmological box simulations, the re-simulation strategy of FLARES means that the creation of the luminosity function (or stellar mass function) is not straightforward. The contribution from any of our re-simulated region needs to be weighted by the appropriate weight for that region. This was explained in §3.3.2.1 with the weighting scheme detailed in §3.2.3.

As described in §4.1.1, we concentrate on a broader definition of a galaxy focusing on only those objects with a combined total of more than 100 gas and star particles, extending the stellar mass function to $\sim 10^{7.5} M_{\odot}$ at $z = 5$. For the luminosity function we set the low brightness cut-off for the selected galaxies to be the 97th percentile of the magnitude computed for 100 gas and star particles, allowing us to probe down to ~ -17 in FUV rest-frame magnitude at $z = 5$. This also means that most of our galaxies have many more than 100 gas and star particles.

4.2.1.2 Luminosity Functions

We plot the dust-attenuated (as described in §4.1.4) UV LF in Figure 4.4 (solid line) along with the intrinsic LF (dashed line). Here the plotted data for FLARES are in bins of width 0.5 magnitudes, with their 1σ Poisson scatter. Also plotted is the UV LF of the 100 cMpc EAGLE Reference simulation box. The luminosity function is extended to brighter galaxies by 2 magnitudes or more at all redshifts, with the Reference volume failing to probe the bright end of the UV LF. It is evident

that at the faint-end the simulations agree. The bin centre and the number density per magnitude for the FLARES galaxies are provided in Appendix 4.B as Table 4.1.

The number density of bright galaxies ($M_{1500} \leq -20$) increases by ~ 2 orders of magnitude going from $10 \rightarrow 5$ in redshift, indicating the rapid assembly of stars in galaxies through time. It can also be seen that the observed LF is slightly lower than the intrinsic LF at luminosities fainter than ~ -20 . The reason for this is the implementation of a birth cloud component for young stellar populations. Studies exploring the impact of birth cloud attenuation have shown that this can reduce the luminosities by ~ 0.3 dex for galaxies in the local Universe (e. g. Trayford et al., 2017). Since the surface density of metals in the faint galaxies is insufficient to produce significant attenuation in the ISM, the choice of birth cloud component is most pronounced in this regime. While in the case of the bright end, the main contribution is from the dust attenuation in the ISM.

It is important to take note that both these regimes can be affected by the choice of initial mass function, the SPS model (see Wilkins et al., 2016a) and the attenuation law. We also do not take into account the contribution of accretion on to super-massive black holes (SMBH) which is expected to dominate over the contribution of star formation at the extreme bright end ($M_{UV} \lesssim -23$ Magnitude at $z \sim 6$, e. g. Glikman et al., 2011; Giallongo et al., 2015; Ono et al., 2018). To give an estimate on the contribution of SMBH to the galaxy luminosity, we perform a simple analysis. The intrinsic bolometric luminosity of the galaxy is compared to the SMBH bolometric luminosity, calculated using

$$L_{\text{BH,bol}} = \eta \frac{dM_{\bullet}}{dt} c^2, \quad (4.6)$$

where dM_{\bullet}/dt is the accretion rate and η is the efficiency, assumed to be 0.1. From this analysis we estimate that the fraction of galaxies where the SMBH bolometric luminosity contributes more than 10% to the total luminosity (intrinsic + SMBH) to be negligible at $M_{1500} < -21$. Below this, the fraction rises to a value of $\sim 10\%$, with a median contribution of $\sim 22\%$ for $z \in [5, 10]$. It is worth noting that these are the bolometric fractions and thus the contribution to the UV can vary widely depending on the obscured nature of the SMBH and the stars. The detailed modelling of SMBH luminosities is the focus of a work in preparation.

A Schechter function (Schechter, 1976) can be used to describe the UV LF (e. g. Bouwens et al., 2015; Finkelstein et al., 2015), characterized by a power law at the faint end with slope α , with an exponential cutoff at the bright end at a characteristic magnitude M^* , with the parameter ϕ^* setting the normalisation of this function. The number density at a given magnitude is then

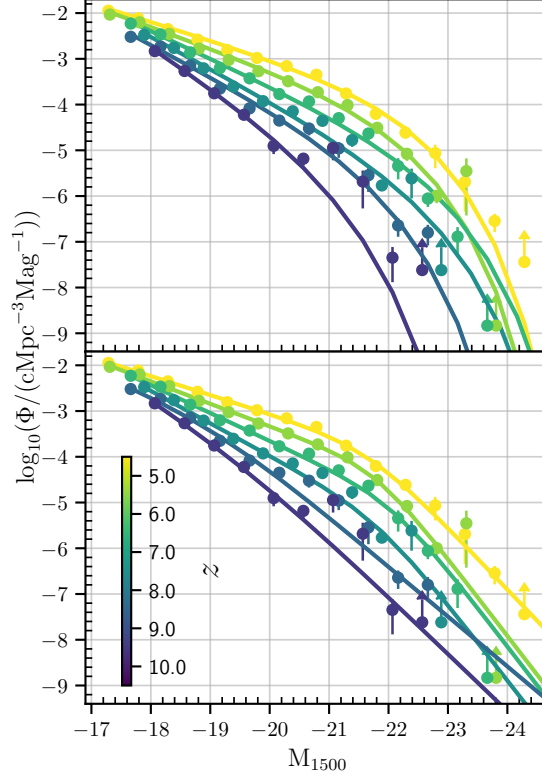


Figure 4.5: Schechter (top) and double power-law (bottom) fits to the FLARES UV LF are plotted as solid lines, while the data is shown as points with $1\text{-}\sigma$ Poisson errors. Bins containing single galaxies are indicated by lower limits.

given by

$$\phi(M) = 0.4 \ln 10 \phi^* 10^{-0.4(M-M^*)(\alpha+1)} e^{-10^{-0.4(M-M^*)}}. \quad (4.7)$$

We calculate the Schechter function parameters of our LFs (see Appendix 4.B for more details of the fitting). The Schechter fits to the UV LF of FLARES galaxies are shown in Figure 4.5 (top panel). We find that the function provides a good fit to the shape of the overall UV LF. The best-fitting Schechter parameters to the UV LF are shown in Table 4.2.

There have also been studies that suggest a double power-law can be used to describe the shape of the UV LF at higher redshifts (e. g. Bowler et al., 2014). We describe the parameterization for a double power-law as follows

$$\phi(M) = \frac{\phi^*}{10^{0.4(M-M^*)(\alpha+1)} + 10^{0.4(M-M^*)(\beta+1)}}, \quad (4.8)$$

where α and β are the faint-end and bright-end slopes, respectively, M^* is the characteristic magnitude between these two power-law regimes, and ϕ^* is the normalisation. The double power-law fit to the binned luminosities is shown in Figure 4.5 (bottom panel). The best-fitting double power-law parameters to the UV LF are also shown in Table 4.2. It can be seen that this also

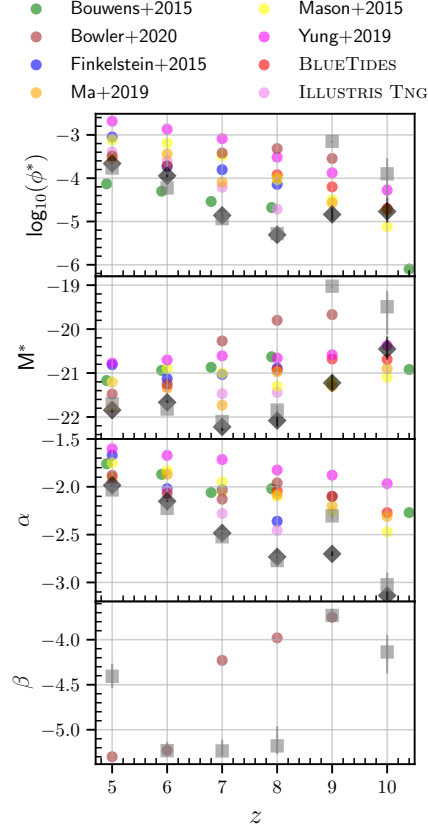


Figure 4.6: Evolution of the parameters of Schechter (black diamonds) and double power-law (grey squares) fits to the FLARES UV LF. The quoted error bars show the 16th – 84th percentile uncertainty obtained from the fit posteriors (see Appendix 4.B for details). Also plotted are the evolution of the Schechter fit parameters from BLUETIDES (Wilkins et al., 2017), Bouwens et al. (2015); Mason et al. (2015); Finkelstein et al. (2015); Ma et al. (2019); Yung et al. (2019a); ILLUSTRIS-TNG (Model-C from Vogelsberger et al., 2020b) as well as the double power-law fit parameters from Bowler et al. (2020).

provides a good fit to the UV LF even though, like the Schechter fit, this parameter form fails to capture the increase in number density around the knee at $z > 8$.

We have already shown in §3.3.2 that the galaxy stellar mass function in FLARES can be described by a double Schechter form. It can be seen in Figure 4.4 that the intrinsic UV LF also has a double Schechter shape, but the observed UV LF does not. It lies much closer to a Schechter or a double power-law shape depending on the redshift. This can be explained by dust attenuation suppressing the intrinsically bright galaxies at the knee and beyond. Also shown is the evolution of the parameters of the Schechter and double power-law fits with redshift in Figure 4.6. We see that for both the fit functions, the value of M^* and α are similar across redshift, with the values generally increasing with increasing redshift for M^* and vice versa for α . The Schechter function shows a smooth evolution in all the parameters while in the case of the double-power law there is a sharp upturn in the parameters ϕ^* , M^* and β . For the purposes of the fitting (also see Appendix 4.B), β was restricted to a lower limit of -5.3, due to the FLARES LF failing to constrain

that parameter. The flattening at $z \sim 7$ can be attributed to this restriction. However, the jump in the parameter space is a consequence of the strong evolution at the bright-end from rapid build up of dust. A similar jump is also seen in the double power-law ‘ β ’ parameter presented in [Bowler et al. \(2020\)](#), albeit at $z = 6 \rightarrow 7$.

We compare the performance of the two functional forms across redshifts by computing the Bayesian Information Criterion (BIC, see [Schwarz, 1978](#); [Liddle, 2007](#), and references therein for further details; also see Appendix 4.B) for the best-fit parameters. A model with a lower BIC is preferred. For this purpose we give the difference between the BIC values of the double power-law from the Schechter best-fit values, which is also quoted in Table 4.2. As can be seen a double power-law function is a much better fit to the UV LF of the FLARES galaxies at all redshifts, except at $z = 10$, where the BIC values are comparable. This could simply be due to the lack of brighter galaxies after the estimated knee of the functions. There are a few explanations in the literature for the emergence of a double power-law shape to the luminosity function at high redshifts. Some studies (e.g. [Bowler et al., 2014, 2020](#)) have suggested that this is due to a lack of evolution in the bright end of the galaxy luminosity function because of the deficit of quenched galaxies at these redshifts. The bright end is very dependent on the dust content as well as star formation of the galaxies, and thus also provides constraints on the recipes of dust modelling and star formation. None of the FLARES regions have galaxies that have moved into the passive regime at $z > 7$, thus it is not surprising that the double power-law performs better at the higher redshifts.

4.2.1.3 Comparison with Observations and Models

In Figure 4.7 the UV LF of FLARES galaxies is compared to observational values from [Bouwens et al. \(2015\)](#); [McLeod et al. \(2015\)](#); [Finkelstein et al. \(2015\)](#); [Bouwens et al. \(2016, 2017\)](#); [Oesch et al. \(2018\)](#); [Atek et al. \(2018\)](#); [Stefanon et al. \(2019\)](#); [Bowler et al. \(2020\)](#).

The UV LF relation of the FLARES galaxies at all redshift is in good agreement within the observational uncertainties. It should also be noted that the uncertainties in the observations gets progressively larger with increasing redshift and some of the number densities at the bright end are upper limits. We slightly over-predict the number density of galaxies at $z = 10$ at the faint-end compared to the observational data. However, the observations at $z = 10$ are limited by the *Hubble Space Telescope*’s capability to detect galaxies, and hence the [Oesch et al. \(2018\)](#) study contain a total of only 9 galaxies. This will change with the imminent launch of *JWST*, which will be able detect a larger sample of galaxies and bring tighter constraints.

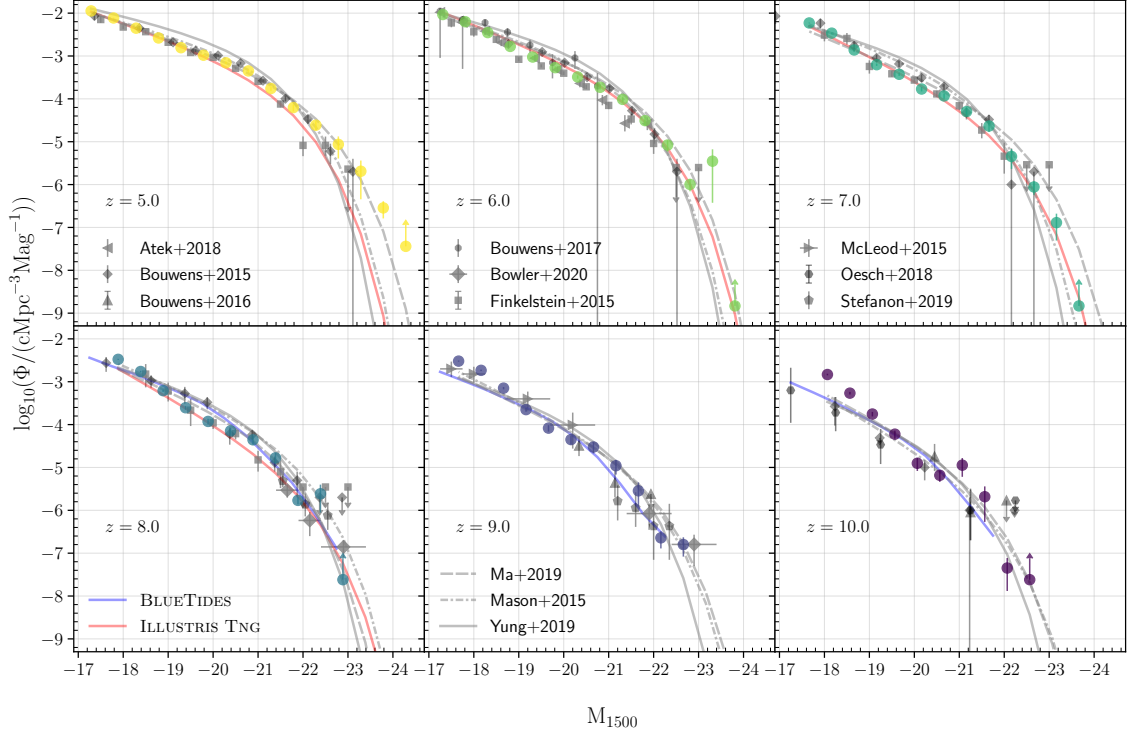


Figure 4.7: UV LF of the FLARES galaxies, represented by the large coloured dots for $z \in [5, 10]$. Error bars denote the Poisson 1σ uncertainties for each bin from the simulated number counts for the dust attenuated UV LF. Observational data from Bouwens et al. (2015); McLeod et al. (2015); Finkelstein et al. (2015); Bouwens et al. (2016, 2017); Oesch et al. (2018); Atek et al. (2018); Stefanon et al. (2019); Bowler et al. (2020) are plotted as well as the binned luminosities from BLUETIDES (Wilkins et al., 2017) and the Schechter fits from Mason et al. (2015); Ma et al. (2019); Yung et al. (2019a); ILLUSTRIS TNG (Model-C from Vogelsberger et al., 2020b) are shown for comparison.

In Figure 4.7, we also plot the binned luminosities from BLUETIDES (Wilkins et al., 2017) and the Schechter function fits from Mason et al. (2015), FIRE-2 (Ma et al., 2019); SANTACRUZ SAM (Yung et al., 2019a); ILLUSTRIS-TNG (Model-C from Vogelsberger et al., 2020b). As can be seen the fit is similar to others from literature, and only starts to diverge slightly at $z \geq 8$, with FLARES having a lower number density at the bright end compared to the Schechter fits from Mason et al. (2015); Ma et al. (2019). Modelling differences across the studies or the larger dynamic range probed by FLARES is a possible explanation for this deviation. With respect to BLUETIDES, a comparison of data have shown us that the most massive galaxies in FLARES are more metal rich by ~ 0.1 dex. This results in increased dust attenuation in FLARES compared to BLUETIDES in , and thus cause differences in the observed UV continuum, attenuation and line luminosity values presented in the next sections. However, a direct comparison to Wilkins et al. (2017, 2020), which also implemented a similar line-of-sight attenuation model, is not possible due to the difference in the modelling approach, namely the implementation of birth cloud attenuation and the dependence on an evolving DTM ratio.

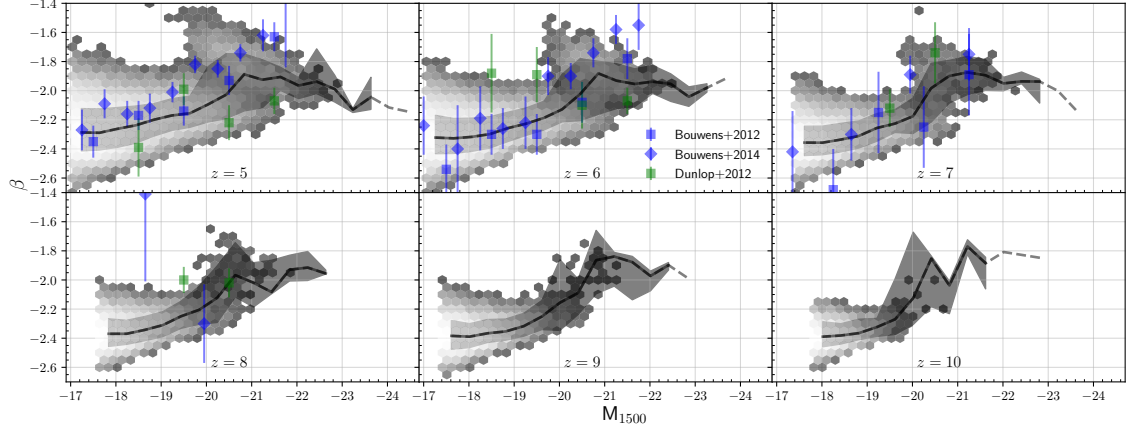


Figure 4.8: Figure shows the UV continuum slope, β , plotted against the UV magnitude for $z \in [5, 10]$. The solid dashed line is the weighted median of the sample, with the shaded region indicating the weighted 84th and 16th percentiles. The hexbin denotes the distribution of our sample. We only plot bins with more than 5 data points. Plotted alongside are observational values from Dunlop et al. (2012); Bouwens et al. (2012, 2014).

In Figure 4.6 we also plot fit parameters from other studies of simulations (Mason et al., 2015; Wilkins et al., 2017; Yung et al., 2019a; Vogelsberger et al., 2020b) as well as observations (Finkelstein et al., 2015; Bowler et al., 2015, 2020). There exists degeneracies between the fit parameters (see Robertson, 2010), and these depend upon the dynamic range and the statistics of the galaxy population. FLARES probes higher density regions, and can therefore better sample the bright end as well as the knee of the function. Thus it is not straightforward to compare fit parameters from different studies.

4.2.2 UV continuum slope (β)

The UV continuum slope β , defined such that $f_\lambda \propto \lambda^\beta$ (Calzetti et al., 1994), is commonly used as a tracer of the stellar continuum attenuation. At high redshifts, the rest-frame UV becomes accessible to optical/near-IR instruments. This has been studied by different groups (e. g. Stanway et al., 2005; Wilkins et al., 2011b; Dunlop et al., 2012; Finkelstein et al., 2012; Bouwens et al., 2014; Bhatawdekar & Conselice, 2021) as it is accessible due to deep near-IR observations using the Wide Field Camera 3 (WFC3) on the *Hubble Space Telescope*. These studies have shown that β is particularly sensitive to the metallicity, age, and especially the dust content within a galaxy, and thus it is a useful quantity to check the reliability of theoretical models. However, it is important to note that β is also strongly dependent upon the modelling assumptions like the choice of the IMF, SPS model, dust modelling and extinction law.

Figure 4.8 plots the value of β against the UV luminosity of the galaxies in FLARES. Observational values of β from Dunlop et al. (2012); Bouwens et al. (2012, 2014) are plotted alongside for

comparison. It should be noted that the observational data shows a lot of scatter and the different datasets do not show the same trends. Our weighted median of β 's match observational values for almost all luminosities. At the bright end, $M_{1500} < -20$ the [Bouwens et al. \(2012, 2014\)](#) data predict much steeper β 's compared to our results, which start to flatten while [Dunlop et al. \(2012\)](#) shows lower values. This could be due to the choice of our extinction curve, a steeper/shallower curve will make for a steeper/shallower relation. The β values are an excellent constraint on the theoretical extinction curves, giving insights into the dust properties within the galaxy (see [Wilkins et al., 2012, 2013; Salim & Narayanan, 2020](#)). We examine a few extinction curves from the literature (namely the Calzetti ([Calzetti et al., 2000](#)), Small Magellanic Cloud ([Pei, 1992](#)) and the curve used in [Narayanan et al. 2018](#)) in Appendix 4.C and plot the effect it has on the UV continuum relation in Figure 4.21 (left panel). We find that the FLARES galaxies prefer a steeper extinction curve similar to the SMC in order to reproduce UV continuum observations. It is interesting to note in this context that [Ma et al. \(2019\)](#) probed the IRX- β relation in the FIRE-2 simulation suite using the radiative transfer code SKIRT ([Baes & Camps, 2015](#)), and obtained a relation which is broadly in agreement with using a simple screen model with the SMC extinction curve.

[Shen et al. \(2020\)](#) showed the relation between $M_{UV} - \beta$ (their Figure 9), obtained from applying SKIRT on the ILLUSTRIS-TNG suite of simulations. Similar to what is seen in Figure 4.8, the β values start to flatten at the bright end. [Wu et al. \(2020\)](#), using the SIMBA simulation suite, capture a similar relation, albeit with a higher normalisation using the [Calzetti et al. \(2000\)](#) extinction law. SIMBA implements a self-consistent dust model, which allows them to infer the dust column density directly and use this in their line-of-sight dust attenuation model. They find that dust attenuation becomes important at $M_{1500} < -18$, while in FLARES it starts only at $M_{1500} \lesssim -21$ at $z = 6$. This extra dust extinction could explain the difference in normalisation seen.

In Figure 4.9 we plot the attenuation in the UV against the UV luminosity, in hexbins coloured by the median β value. The value of the attenuation provides insight into the amount of obscured star formation that is going on in galaxies (also see §4.3). Overall, brighter galaxies suffer more attenuation, which is expected as they have had more time to produce stars thus enriching the ISM. We can also see that there is a sudden increase in the UV-attenuation for galaxies brighter than -20 magnitude, pointing towards the rapid build-up of dusty galaxies in this regime. The figure also shows that many of the galaxies at the bright end are not the most attenuated ones. These are the galaxies that have enjoyed a recent burst of star formation and have not had time to enrich the ISM with dust. Another alternative is stellar migration (see [Furlong et al., 2015](#)),

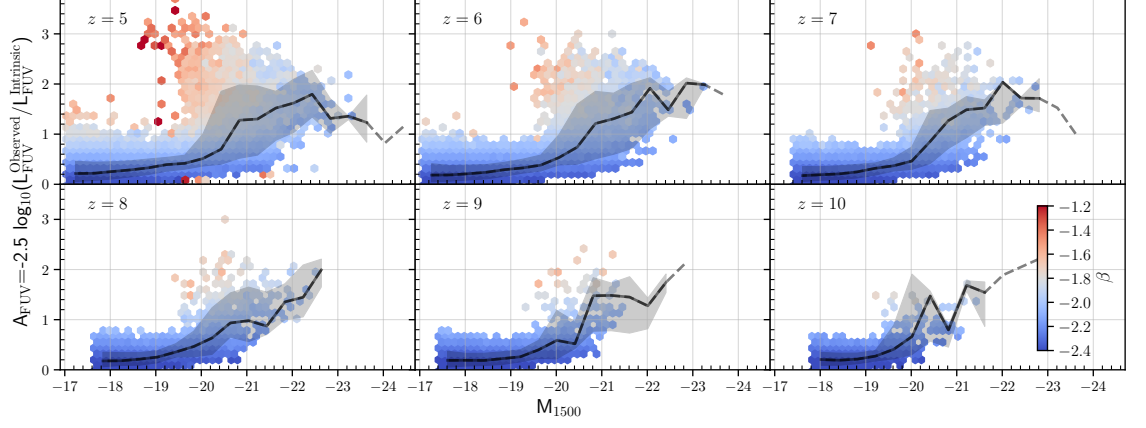


Figure 4.9: Figure shows the attenuation in the FUV against the observed UV magnitude for $z \in [5, 10]$. The solid and dashed black line is the weighted median of the sample, with the shaded region indicating the weighted 84th and 16th percentiles. The dashed line is for bins that have less than 3 data points. The hexbin denotes the distribution of our sample, coloured by the median β value in the hexbin.

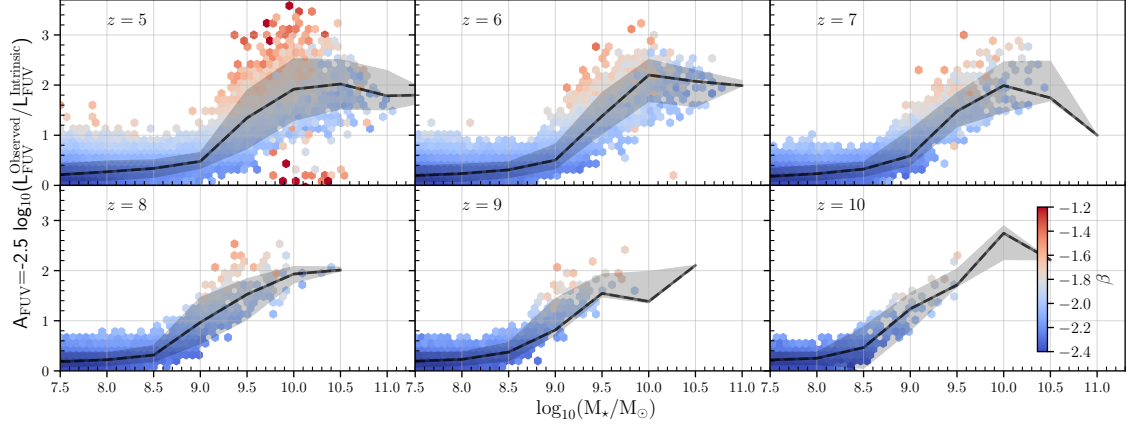


Figure 4.10: Figure shows the attenuation in the FUV against the galaxy stellar mass for $z \in [5, 10]$. The solid and dashed black line is the weighted median of the sample, with the shaded region indicating the weighted 84th and 16th percentiles. The hexbin denotes the distribution of our sample, coloured by the median β value in the hexbin.

with some stars moving radially outwards, thus subject to reduced dust attenuation depending on the viewing angle or geometry. The observed UV LF being better fit by double power-law at these high-redshift (seen in observations, for e. g. [Bowler et al., 2017, 2020](#); [Shibuya et al., 2021](#)) also points towards a decrease in the dust attenuation at the bright and massive end. Some recent *ALMA* studies at high redshift (e. g. [Bowler et al., 2018](#)) have also found galaxies having a heavily dust-obscured and an unobscured component. The variation of dust attenuation within a galaxy as well as the viewing angle will be explored in a future work. [Ma et al. \(2020\)](#), using the FIRE-2 simulation, studied the escape fraction of ionising photons across different resolutions, and found that the lowest resolution run had a lower escape fraction compared to the higher resolutions. In a future study we plan to explore the effect of dust attenuation with resolution on our dust model.

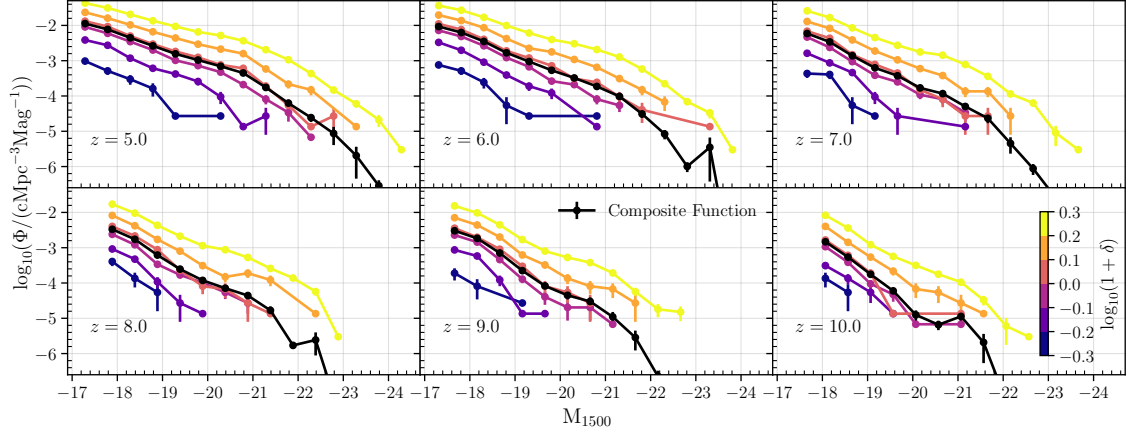


Figure 4.11: The FLARES UV LF for $z \in [5, 10]$ split by binned log-overdensity. Error bars denote the Poisson 1σ uncertainties for each bin from the simulated number counts.

In Figure 4.20 we plot the attenuation as a function of intrinsic FUV luminosity. This provides more insights into the features seen in Figure 4.9; in general, intrinsically brighter galaxies are more attenuated. A comparison also reveals that many of the intrinsically bright galaxies, since they are dusty, are not the brightest galaxies observed in the UV. The relations presented above are also in agreement with the $A_{UV} - M_{\star}$ and the $A_{UV} - \beta$ relations presented in Shen et al. (2020) (their Figures 10 and 11) at $z \leq 6$.

We also plot the attenuation as a function of galaxy stellar mass in Figure 4.10. Features similar to the plots described earlier are seen here as well, with a flattening of the relation at the low mass end ($\lesssim 10^{8.5} M_{\odot}$), and rapid steepening afterwards. As seen in local observations our values do not exhibit a large scatter at the low mass end. This scatter at low redshift can be explained by varying dust content and star-dust geometries of the galaxies. High resolution simulations such as FIRE-2 (see Ma et al., 2019) also see a flattening of the FUV attenuation at the low mass end, with more scatter, possibly due to the low number galaxies produced at the massive end.

We have examined the few galaxies at $z = 5$ that have very low attenuation, but have high β values (also seen in Figure 4.10). They also are intrinsically very bright (see Figure 4.20). These are galaxies that are identified to be in the passive regime, whose specific star formation rate was calculated to be $\lesssim 1/(3 \times H(z))$, where $H(z)$ is the Hubble constant at $z = 5$. We will be studying this population in more detail in a future work.

4.2.3 Effect of environment

FLARES probes galaxies that reside in a wide range of environments allowing us to analyse the effect environment has on their observed properties. In Figure 4.11 we look at how the UV LF varies as a function of overdensity for $z \in [5, 10]$. Here we have plotted the UV LF in 6 bins of $\log_{10}(1 + \delta)$, where δ is the overdensity. As expected the number density of galaxies increases with increasing overdensity and the brighter galaxies reside predominantly in denser environments. Similar behaviour has been seen in measurements of the UV LF in high-redshift galaxy protoclusters (Ito et al., 2020). The normalisation shows a variation of ~ 2 dex from the lowest to the highest density environment probed in FLARES, much greater than the 0.5 dex variation in density itself. The composite distribution function closely follows that of mild overdensity, $\log_{10}(1 + \delta) \in 0 - 0.1$, with the contribution to the bright end coming only from the densest environments.

As can be seen from Figure 4.11 the shape of the luminosity function is similar across various environments with no significant variation in the knee of the function. There is a hint of a double Schechter shape, being strongest in intermediate to lower density environments at high redshift. This could be due to the different assembly histories of galaxies driven by the environment. The effect of environment on assembly history as well as on astronomical surveys will be probed in a future work.

We have also looked at the UV continuum slopes as well as the attenuation in the far-UV as a function of environment similar to the method described above. We find no dependence on overdensity for these galaxy properties.

4.2.4 Line Luminosities and Equivalent Widths

In this section we will present some of the nebular emission line properties and compare them to some of the available observational constraints.

We present predictions for 6 prominent nebular lines or doublets in the UV in Figure 4.12. The top panel shows the evolution of the line luminosity function with redshift, for $z \in [5, 10]$. The overall shape of the function is similar to the UV luminosity function of galaxies and can be approximated by a Schechter function at these redshifts. The LF of the lines evolves with redshift, with almost 3 dex in value near the knee of the function. We also present predictions for the evolution of the weighted median equivalent widths of these lines as a function of stellar mass (middle panel) and far-UV luminosity (bottom panel) with redshift in Figure 4.12. For

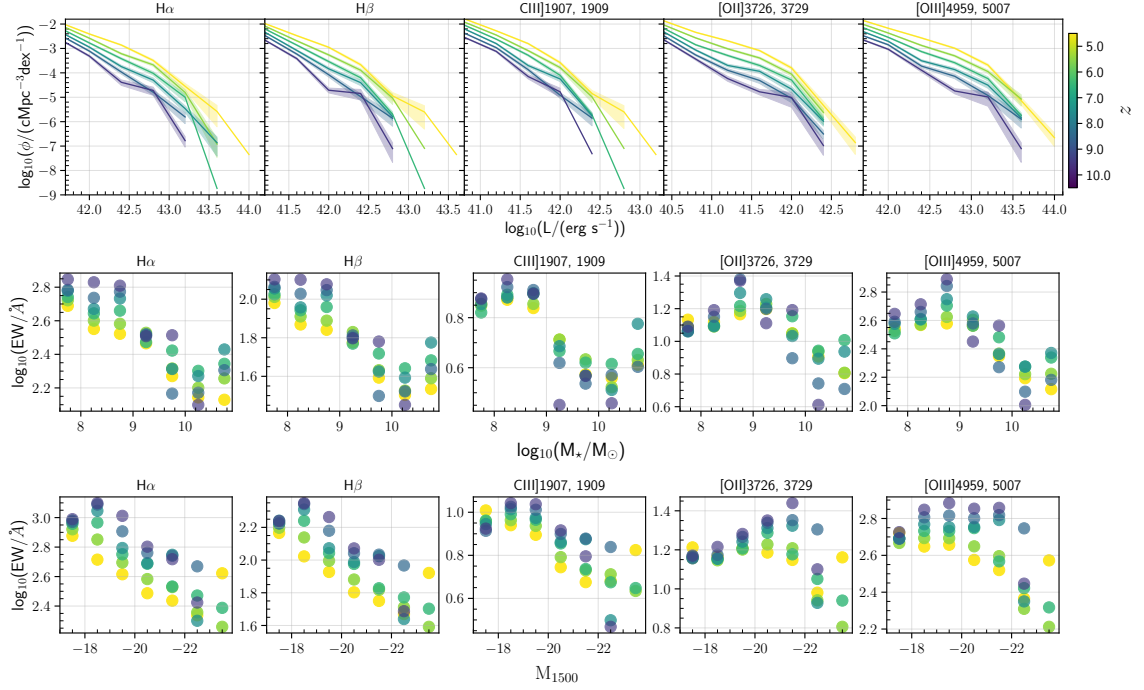


Figure 4.12: Predictions for the properties of 6 prominent UV and optical lines in FLARES for $z \in [5, 10]$. The colour bars for the different redshifts are shown in the rightmost panel. In the top panel we show the dust-attenuated luminosity functions for each line, with the shaded region representing the 1σ Poisson uncertainties. Middle panel shows the evolution of the weighted median equivalent widths of these lines in stellar mass bins. Bottom panel shows the weighted median equivalent widths as a function of FUV luminosity.

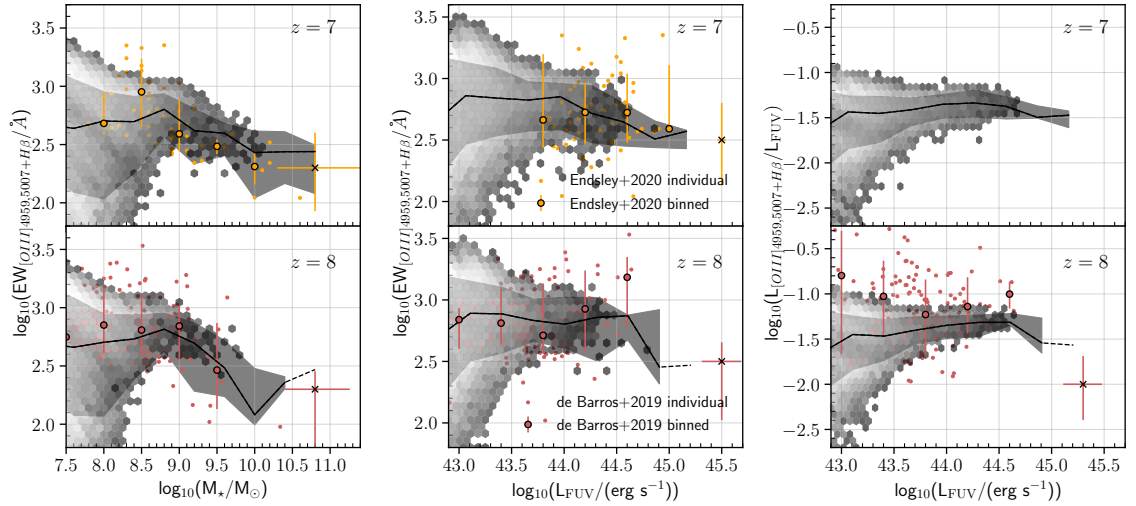


Figure 4.13: Left: Predicted distribution of combined $H\beta$ and $[OIII]\lambda 4959,5007$ equivalent widths and stellar masses for FLARES galaxies at $z \sim 7, 8$. Middle: Predicted distribution of combined $H\beta$ and $[OIII]\lambda 4959,5007$ equivalent widths to the far-UV luminosity of FLARES galaxies at $z \sim 7, 8$. Right: Predicted distribution of the $H\beta$ and $[OIII]\lambda 4959,5007$ line luminosities to the far-UV luminosity and far-UV luminosities of FLARES galaxies at $z \sim 7, 8$. The solid line is the weighted median of the sample, with the shaded region indicating the weighted 84th and 16th percentiles. The hexbin denotes the distribution of our sample, only plotted are bins with more than 5 data points. The small circles show the individual measurements from [De Barros et al. \(2019\)](#); [Endsley et al. \(2021\)](#) while the large points denote the median value in bins of stellar mass and far-UV luminosities respectively. The errorbars centered on the cross shown at the bottom-right gives the median errors on the observational data.

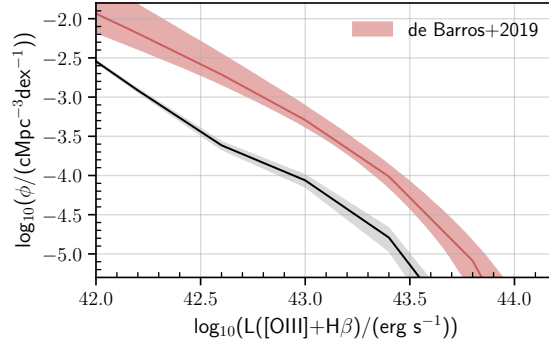


Figure 4.14: The De Barros et al. (2019) and predicted combined $H\beta$ and $[OIII]\lambda 4959,5007$ line luminosity function of FLARES galaxies at $z \sim 8$.

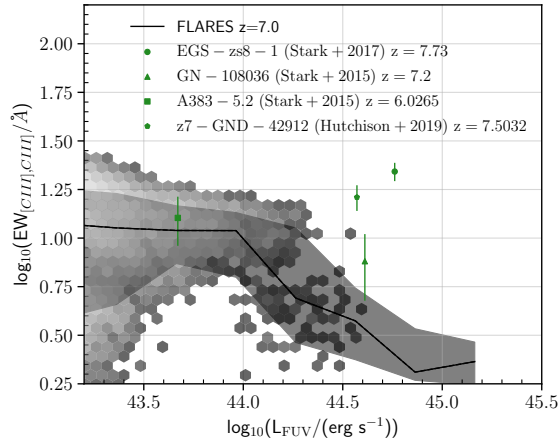


Figure 4.15: Predicted $[CIII]\lambda 1907, \lambda 1909$ line equivalent widths of FLARES galaxies at $z \sim 7$. The solid line is the weighted median of the sample, with the shaded region indicating the weighted 84th and 16th percentiles. The hexbin denotes the distribution of our sample, only plotted are bins with more than 5 data points. Plotted alongside are observational values from Stark et al. (2015, 2017); Hutchison et al. (2019).

galaxies with similar stellar mass the equivalent width mostly increases with increasing redshift, indicating that they have harder ionising photons from their younger stellar population with more massive stars. There is also the effect of metallicity on the line width, causing them to drop quickly at higher stellar masses in case of the hydrogen recombination lines, while the other lines peak around $10^9 M_\odot$ and then fall rapidly. In case of the far-UV, the relationship with metallicity is not correlated in the same way as stellar mass and hence interpretation is harder. But in most cases this also shows increasing equivalent widths at higher redshifts for fixed far-UV luminosity. This behaviour is in agreement with that seen from the BLUE TIDES simulation presented in Wilkins et al. (2020).

Both De Barros et al. (2019); Endsley et al. (2021) have combined broadband photometry from *Hubble* and *Spitzer* observations to constrain the prominent $H\beta$ and $[OIII]\lambda 4959,5007$ lines at $z \sim 7, 8$. In Figure 4.13 we plot the combined values of $[OIII]\lambda 4959,5007$ and $H\beta$ line luminosities as well as the equivalent widths (EWs) of FLARES galaxies at $z = 7, 8$ against these observa-

tional data sets. As can be seen from the figure, in the case of the equivalent width measurements plotted against the stellar mass (left panel) or FUV luminosity (middle panel), the weighted median closely follows the observations. However, it should be noted that our modelling does fail to reproduce some of the larger values of the EW measurements. In case of the line luminosity normalised by the far-UV luminosity (right panel), FLARES lies ~ 0.3 dex below the observational data from [De Barros et al. \(2019\)](#). We also compare the $[\text{OIII}]\lambda_{4959,5007}$ luminosity function as predicted by [De Barros et al. \(2019\)](#) at $z = 8$ to FLARES in Figure 4.14. Our result is offset by ≈ 0.6 to lower number densities or by ≈ 0.4 to lower luminosities. The cause of this offset could be due to the relation used by [De Barros et al. \(2019\)](#) to convert the observed far-UV LF to a line luminosity LF. A similar feature is also seen in the $z = 8$ $[\text{OIII}]\lambda_{4959,5007}$ luminosity function from the ILLUSTRIS-TNG simulation presented in [Shen et al. \(2020\)](#) (their Figure 5), with marginal consistency at the bright end (> 43.5 erg/s). [Wilkins et al. \(2020\)](#) also show an underprediction of the luminosity function at $z = 8$.

We also show the predicted $[\text{CIII}]\lambda_{1907,\lambda_{1909}}$ line equivalent widths of FLARES galaxies at $z \sim 7$ against observations from [Stark et al. \(2015, 2017\)](#); [Hutchison et al. \(2019\)](#) in the redshift range of 6 – 8 in Figure 4.15. A similar feature is seen here as well where we underpredict some high-EW measurements at the most luminous end. An explanation of this discrepancy could be due to the assumptions in the nebular emission modelling like the nebular hydrogen density or ionisation parameter (using a distribution of reference ionisation parameter values instead of a single one) as well as contributions from AGN which we have not considered in this work (see Section 3.4 in [Wilkins et al., 2020](#), for more details). Future direct emission line measurements from *JWST* and other facilities will help to constrain this observational space and thus better understand this discrepancy.

4.3 SFR distribution functions

The instantaneous SFR distribution function of the FLARES galaxies was already presented in §3.3.3, which followed a double Schechter form and provided a good match to the observed values. In this section we look at the relative contribution of the obscured and unobscured/uncorrected star formation rate in FLARES. We compute the fraction of obscured star formation or infrared star formation rate, f_{obsc} going on in any given galaxy by using the attenuation in the far-UV, A_{FUV} . It is computed as

$$f_{\text{obsc}} = 1 - \frac{L_{\text{FUV}}^{\text{Observed}}}{L_{\text{FUV}}^{\text{Intrinsic}}} = 1 - 10^{-A_{\text{FUV}}/2.5}, \quad (4.9)$$

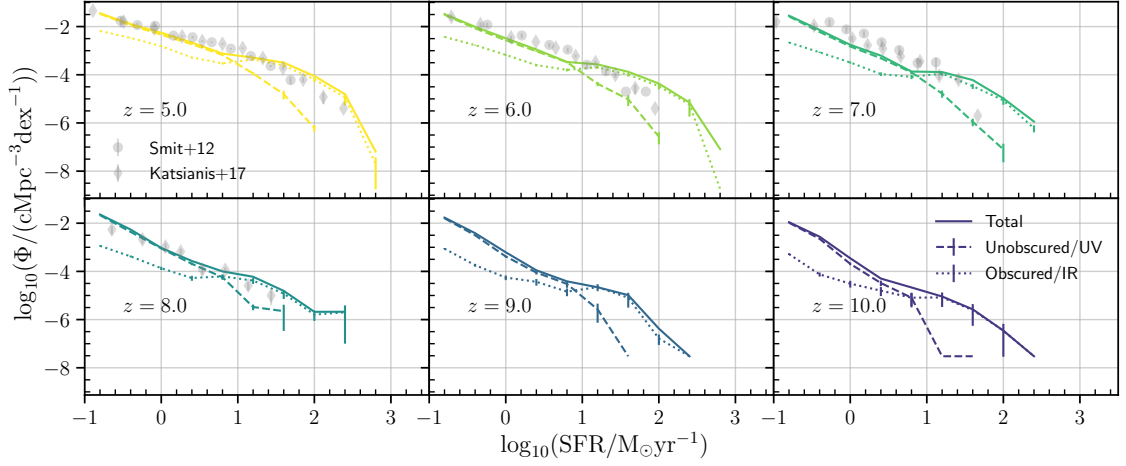


Figure 4.16: FLARES composite galaxy total (solid), obscured (dotted) and unobscured (dashed) star formation rate function for $z \in [5, 10]$. The $1-\sigma$ Poisson uncertainties for the obscured and unobscured star formation rate function are also plotted. For comparison the dust-corrected SFRF from Smit et al. (2012); Katsianis et al. (2017) is also shown.

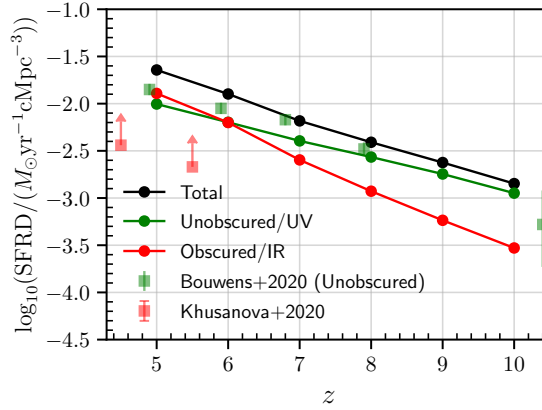


Figure 4.17: FLARES composite galaxy total (black circles), obscured (red circles) and unobscured (green circles) star formation rate density for $z \in [5, 10]$. For comparison the uncorrected SFRD or SFRD_{UV} from Bouwens et al. (2020) (obtained from UV luminosity scaling relations) and SFRD_{IR} from Khusanova et al. (2020) (which are lower limits) is also shown.

with $f_{\text{unobsc}} = 1 - f_{\text{obsc}}$ the fraction of unobscured star formation rate. Using this prescription, the rate of obscured (infrared) and unobscured (far-UV) star formation rate are $f_{\text{obsc}} \times \text{SFR}$ and $f_{\text{unobsc}} \times \text{SFR}$, respectively. This would differ slightly from the observed calibration, where the obscured and unobscured SFRs are obtained by combining the total IR and observed UV luminosities with a theoretically motivated calibration (e.g. Kennicutt Jr & Evans II, 2012). We use the SFR of a galaxy averaged over the star particles that were formed in the last 100 Myr. These would closely resemble SFRs inferred observationally from the UV/IR, rather than ones that were obtained by emission line calibrations.

Figure 4.16 shows the total, obscured and unobscured SFR distribution function for the FLARES galaxies in $z \in [5, 10]$. We also plot the dust-corrected SFR function from Smit et al. (2012);

Katsianis et al. (2017) for comparison. The dust corrections are done using the IRX- β relation established by Meurer et al. (1999). This can be uncertain for highly star-forming systems and possibly underestimated (Katsianis et al., 2017). As can be seen, obscured star formation dominates the contribution to the total at SFRs $\gtrsim 10 M_{\odot}/\text{yr}$, indicating the rapid build up of dust in these extreme star forming galaxies. This directly reflects what is seen in the FUV attenuation that is presented in Figures 4.9, 4.20, and 4.10, where there is a rapid increase in the attenuation when moving to the very-bright/massive end of the distribution.

We also look at the evolution of the total (black), obscured (red) and unobscured (green) star formation rate density (SFRD) in Figure 4.17 for galaxies with $\text{SFR} \geq 0.1 M_{\odot}/\text{yr}$. Even though the bright end is dominated by obscured star formation at all redshifts, we find that the contribution to the total SFRD is mainly coming from unobscured star formation that takes place in low mass galaxies, or specifically from galaxies below the knee of the SFR function. The contribution of obscured star formation is $\sim 40\%$ at $z = 7$ and becomes almost equal at $z \sim 6$. This is similar to the fraction of obscured star formation found in recent observational surveys with *ALMA* (e. g. Khusanova et al., 2020), where they predict the SFRD_{IR} to possibly cross the SFRD_{FUV} at $z > 5$. Bouwens et al. (2020) also see a transition of the SFR density being primarily unobscured at $z > 5$ and obscured at $z < 5$. We plot these measurements for comparison in Figure 4.17.

4.4 Conclusions

We have presented the photometric results from the FLARE simulations, a suite of zoom simulations run using the EAGLE (Schaye et al., 2015; Crain et al., 2015) simulation model probing a wide range of overdensities in the Epoch of Reionisation ($z \geq 5$). The wide range of overdensities sampled from a large periodic volume allows us to probe brighter and more massive galaxies in the EoR. Using a simple line-of-sight dust extinction model we retrieve the photometric properties of the galaxies in the simulation. Our main findings are as follows:

- The FLARES UV LF provides an excellent match to current observations of high-redshift galaxies. The UV LF exhibits a double power-law form at all redshifts with the Schechter form being comparable at $z = 10$ from BIC. The number density of bright objects at the knee of the function increases by almost 2 orders of magnitude from the lowest to the highest density environment probed in FLARES. The normalisation of the UV LF is strongly dependent on the environment, with the shape being affected to a lesser extent.
- The relationship between the UV continuum slope, β and M_{1500} of the FLARES galaxies are

in very good agreement with the observations. We find a flattening of the relation at the bright-end.

- The attenuation in the far-UV also shows a linear relationship with the observed as well as the intrinsic UV luminosity. There is a sudden increase in the UV-attenuation for galaxies brighter than -20 magnitude, pointing towards the rapid build-up of dusty galaxies in this regime. The brightest objects in the UV are not the most attenuated.
- We find good agreement of observed line luminosity and equivalent width relationship of the combined $[\text{OIII}]\lambda_{4959,5007}$ and $\text{H}\beta$ lines as well as the $\text{CIII}]\lambda_{1907}, [\text{CIII}]\lambda_{1909}$ line equivalent widths.
- The star formation in galaxies with a $\text{SFR} \gtrsim 10 \text{ M}_{\odot}/\text{yr}$ is predominantly obscured and vice versa below that for FLARES galaxies in $z \in [5, 10]$. Dust obscured star formation (for galaxies with $\text{SFR} \geq 0.1 \text{ M}_{\odot}/\text{yr}$) makes a significant contribution at these high redshifts reaching $\sim 40\%$ at $z = 7$, and starts dominating below $z \sim 6$.

Future observations from *Webb*, *Euclid* and the *Roman Space Telescope* will provide further constraints on the photometric properties of these high redshift galaxies. Complimentary observations in the far-IR by *ALMA* will also be instrumental in providing additional constraints on the nebular emission characteristics.

4.A Calibrating Dust Attenuation

As noted in §4.1.4 we model the attenuation by dust on a star particle by star particle basis using the integrated line-of-sight surface density of metals as a proxy for dust attenuation. In this simple model we have a two free parameter κ_{BC} and κ_{ISM} which encapsulates the properties of dust such as the average grain size, shape, composition in the birth clouds and in the ISM respectively. In case of birth clouds, κ_{BC} also incorporates the dust-to-metal ratio, which is assumed to scale linearly with the metallicity of the stellar particle. We calibrate these two parameters by comparing to observations of the UV LF at $z = 5$ from Bouwens et al. (2015), UV-continuum slope (β) at $z = 5$ from Bouwens et al. (2012, 2014) as well as the line luminosity and the EW relation of $[\text{OIII}]\lambda_{4959,5007} + \text{H}\beta$ at $z = 8$ from De Barros et al. (2019). As explained in § 4.1.4 we use a simple grid search to calibrate these parameters against these observations. For that purpose we generate a range of values from $[0.001, 2]$ for the parameter κ_{BC} . The required photometric properties¹ are generated from κ_{ISM} values in the range $(0, 1]$. The κ_{ISM} value corresponding to

¹ Photometric properties are generated using the code SYNTHOBS: <https://github.com/stephenmwilkins/SynthObs>

a given κ_{BC} value is chosen to best match the UV LF from [Bouwens et al. \(2015\)](#) at $z = 5$. We generate the UV LF of the FLARES galaxies for a given $(\kappa_{\text{BC},i}, \kappa_{\text{ISM},j})$ pair, where ‘ i ’ and ‘ j ’ corresponds to a position on the grid for these parameters. The simulated and the observed UV LFs are then compared, using a chi-squared analysis to choose the best fit value of κ_{ISM} for the corresponding $\kappa_{\text{BC},i}$. In order to select the combination of these two values that was used in this study, we compare the simulated $M_{\text{UV}} - \beta$ relation at $z = 5$ against [Bouwens et al. \(2012, 2014\)](#), shown in Figure 4.18. As can be seen from the figure, this parameter space prefers a higher value of κ_{BC} for a better fit with the observational data. We tried values of $\kappa_{\text{BC}} > 2$ and found that the median β values have started to converge for those choices. In order to get a measure on the upper limit of κ_{BC} , we compare the simulated outputs of the line luminosity and the EW relation of $[\text{OIII}]\lambda 4959,5007 + \text{H}\beta$ at $z = 8$ from our range of κ_{BC} choices, to the results from [De Barros et al. \(2019\)](#) in Figure 4.19. As can be deduced from the figure, in this case κ_{BC} prefers smaller values. In order to incorporate the impact of both these observations, we choose a value of $\kappa_{\text{BC}} = 1$. The corresponding value of κ_{ISM} is 0.0795, for this choice. Another caveat is that by fixing these values we assume there is no evolution in the general properties of the dust grains with redshift or among different galaxies.

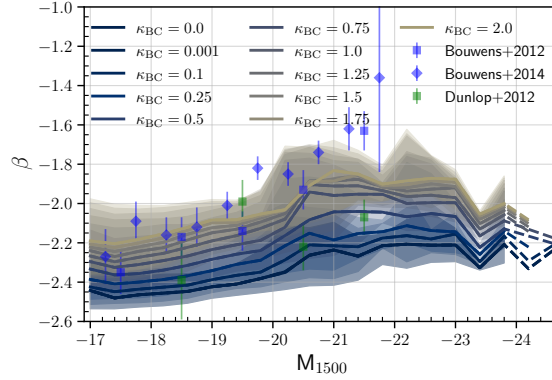


Figure 4.18: UV continuum slope β for different values of κ_{BC} at $z = 5$. Also plotted are the observational data from [Dunlop et al. \(2012\)](#); [Bouwens et al. \(2012, 2014\)](#).

Also presented is the relationship between the intrinsic luminosity of the galaxy and the attenuation in the far-UV in Figure 4.20. The hexbins are coloured by their median UV continuum values (β) with the solid black line showing the weighted median and the shaded region around it representing the 84 and 16 percentiles of the data. The shape is quite similar to Figure 4.9 where the attenuation is plotted against the UV luminosity, and the median increases with the intrinsic luminosity and starts flattening afterwards. Also can be seen at $z = 5$ is a few of the passive galaxies that have high luminosity and high β but lower attenuation.

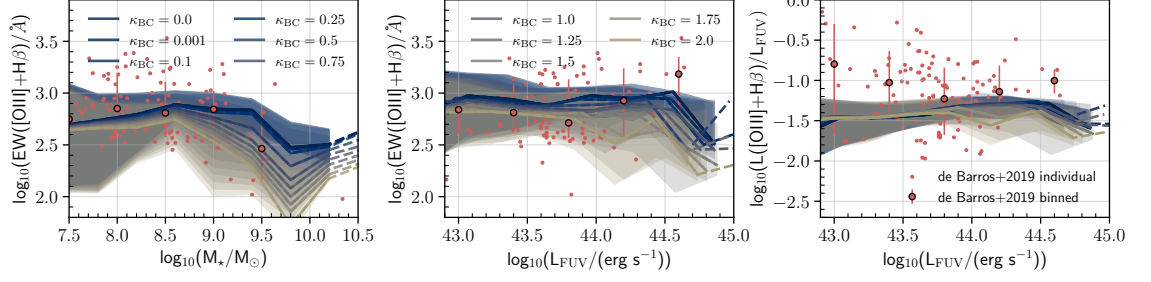


Figure 4.19: Same as Figure 4.13, now showing the line luminosity and equivalent width for different values of κ_{BC} . The small red circles show the individual measurements from De Barros et al. (2019) while the large points denote the median value in bins of stellar mass and far-UV luminosities respectively.

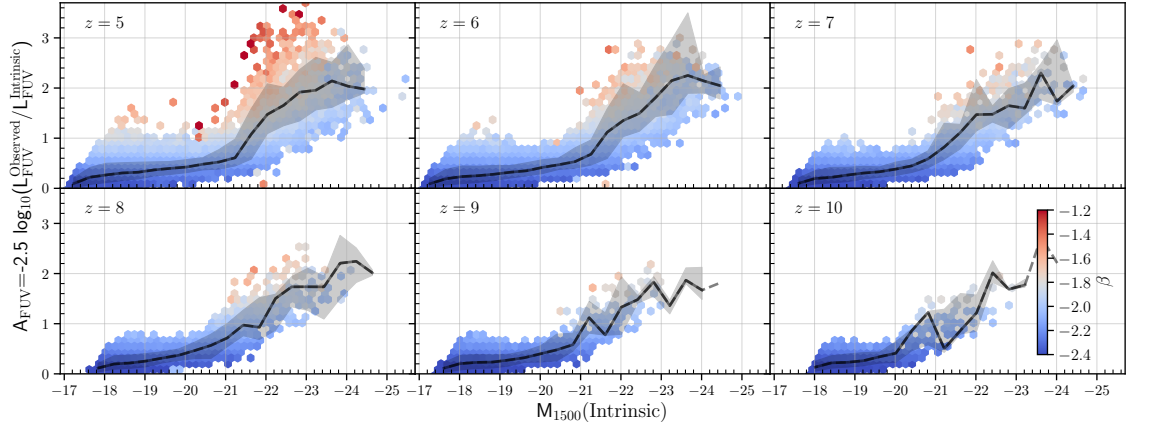


Figure 4.20: Same as Figure 4.9 and 4.10, now showing the attenuation as a function of intrinsic UV luminosity.

4.B UV LF

For deriving the Schechter and double power-law fit parameters for the UV LF, we calculate the likelihood that the number of observed galaxies in a given magnitude bin is equal to that for an assumed value of the function parameters. This calculation is performed in bins of separation $\Delta M = 0.5$ magnitude, ranging from from our completeness limit at the faint-end to enclose all our galaxies above this limit. Bins containing less than 5 galaxies were not considered while fitting. The bin centre and the number density of galaxies per magnitude is provided in Table 4.1. We use the code `FitDF`² a Python module for fitting arbitrary distribution functions. `FitDF` uses `EMCEE`, a Python implementation of the affine-invariant ensemble sampler for Markov chain Monte Carlo (MCMC) described in Foreman-Mackey et al. (2013). The likelihood function is modelled as a Gaussian distribution of the following form

$$\ln(\mathcal{L}) = -\frac{1}{2} \sum_i \left[\frac{(n_{i,\text{obs}} - n_{i,\text{exp}})^2}{\sigma_i^2} + \log(\sigma_i^2) \right], \quad (4.10)$$

² <https://github.com/flaresimulations/fitDF>

M_{1500}	$\phi / (\text{cMpc}^{-3} \text{ Mag}^{-1})$	M_{1500}	$\phi / (\text{cMpc}^{-3} \text{ Mag}^{-1})$	M_{1500}	$\phi / (\text{cMpc}^{-3} \text{ Mag}^{-1})$
$z = 5$		$z = 6$		$z = 7$	
-24.286	$(3.620 \pm 3.620) \times 10^{-8}$	-23.810	$(1.473 \pm 1.473) \times 10^{-9}$	-23.662	$(1.473 \pm 1.473) \times 10^{-9}$
-23.786	$(2.857 \pm 1.235) \times 10^{-7}$	-23.310	$(3.513 \pm 3.137) \times 10^{-6}$	-23.162	$(1.295 \pm 0.801) \times 10^{-7}$
-23.286	$(2.047 \pm 1.593) \times 10^{-6}$	-22.810	$(1.008 \pm 0.311) \times 10^{-6}$	-22.662	$(8.790 \pm 3.015) \times 10^{-7}$
-22.786	$(8.674 \pm 4.616) \times 10^{-6}$	-22.310	$(8.369 \pm 2.476) \times 10^{-6}$	-22.162	$(4.532 \pm 2.214) \times 10^{-6}$
-22.286	$(2.433 \pm 0.691) \times 10^{-5}$	-21.810	$(3.103 \pm 0.726) \times 10^{-5}$	-21.662	$(2.326 \pm 0.632) \times 10^{-5}$
-21.786	$(6.266 \pm 1.186) \times 10^{-5}$	-21.310	$(9.729 \pm 1.518) \times 10^{-5}$	-21.162	$(5.044 \pm 1.114) \times 10^{-5}$
-21.286	$(1.745 \pm 0.201) \times 10^{-4}$	-20.810	$(1.864 \pm 0.210) \times 10^{-4}$	-20.662	$(1.168 \pm 0.164) \times 10^{-4}$
-20.786	$(4.484 \pm 0.339) \times 10^{-4}$	-20.310	$(3.242 \pm 0.289) \times 10^{-4}$	-20.162	$(1.698 \pm 0.205) \times 10^{-4}$
-20.286	$(7.127 \pm 0.438) \times 10^{-4}$	-19.810	$(5.348 \pm 0.373) \times 10^{-4}$	-19.662	$(3.745 \pm 0.320) \times 10^{-4}$
-19.786	$(1.043 \pm 0.053) \times 10^{-3}$	-19.310	$(9.458 \pm 0.517) \times 10^{-4}$	-19.162	$(6.270 \pm 0.406) \times 10^{-4}$
-19.286	$(1.562 \pm 0.066) \times 10^{-3}$	-18.810	$(1.675 \pm 0.069) \times 10^{-3}$	-18.662	$(1.381 \pm 0.062) \times 10^{-3}$
-18.786	$(2.634 \pm 0.087) \times 10^{-3}$	-18.310	$(3.515 \pm 0.101) \times 10^{-3}$	-18.162	$(3.411 \pm 0.099) \times 10^{-3}$
-18.286	$(4.458 \pm 0.115) \times 10^{-3}$	-17.810	$(6.299 \pm 0.137) \times 10^{-3}$	-17.662	$(5.898 \pm 0.133) \times 10^{-3}$
-17.786	$(7.703 \pm 0.152) \times 10^{-3}$	-17.310	$(9.274 \pm 0.167) \times 10^{-3}$	–	–
-17.286	$(1.126 \pm 0.018) \times 10^{-2}$	–	–	–	–
$z = 8$		$z = 9$		$z = 10$	
-22.888	$(2.407 \pm 2.407) \times 10^{-8}$	-22.662	$(1.588 \pm 0.758) \times 10^{-7}$	-22.567	$(2.407 \pm 2.407) \times 10^{-8}$
-22.388	$(2.429 \pm 1.545) \times 10^{-6}$	-22.162	$(2.279 \pm 0.990) \times 10^{-7}$	-22.067	$(4.503 \pm 3.192) \times 10^{-8}$
-21.888	$(1.706 \pm 0.328) \times 10^{-6}$	-21.662	$(2.852 \pm 1.624) \times 10^{-6}$	-21.567	$(2.075 \pm 1.538) \times 10^{-7}$
-21.388	$(1.675 \pm 0.484) \times 10^{-5}$	-21.162	$(1.098 \pm 0.414) \times 10^{-5}$	-21.067	$(1.130 \pm 0.526) \times 10^{-5}$
-20.888	$(4.410 \pm 1.002) \times 10^{-5}$	-20.662	$(3.000 \pm 0.880) \times 10^{-5}$	-20.567	$(6.563 \pm 1.951) \times 10^{-6}$
-20.388	$(7.125 \pm 1.379) \times 10^{-5}$	-20.162	$(4.470 \pm 1.041) \times 10^{-5}$	-20.067	$(1.251 \pm 0.423) \times 10^{-5}$
-19.888	$(1.186 \pm 0.178) \times 10^{-4}$	-19.662	$(8.275 \pm 1.420) \times 10^{-5}$	-19.567	$(5.984 \pm 1.237) \times 10^{-5}$
-19.388	$(2.473 \pm 0.254) \times 10^{-4}$	-19.162	$(2.236 \pm 0.244) \times 10^{-4}$	-19.067	$(1.764 \pm 0.214) \times 10^{-4}$
-18.888	$(6.183 \pm 0.409) \times 10^{-4}$	-18.662	$(7.084 \pm 0.444) \times 10^{-4}$	-18.567	$(5.418 \pm 0.387) \times 10^{-4}$
-18.388	$(1.732 \pm 0.070) \times 10^{-3}$	-18.162	$(1.844 \pm 0.073) \times 10^{-3}$	-18.067	$(1.473 \pm 0.064) \times 10^{-3}$
-17.888	$(3.329 \pm 0.098) \times 10^{-3}$	-17.662	$(3.027 \pm 0.094) \times 10^{-3}$	–	–

Table 4.1: Binned UV LF values for the FLARES galaxies. Also quoted is the weighted $1-\sigma$ Poisson uncertainty for the number density within each luminosity bin.

z	M^*/Mag	$\log_{10}(\phi^*/(\text{Mpc}^{-3} \text{ Mag}^{-1}))$	α	β	ΔBIC
5	$-21.844^{+0.041}_{-0.042}$	$-3.662^{+0.025}_{-0.025}$	$-1.984^{+0.006}_{-0.006}$	—	66.440
	$-21.699^{+0.035}_{-0.042}$	$-3.766^{+0.022}_{-0.026}$	$-2.033^{+0.006}_{-0.007}$	$-4.406^{+0.139}_{-0.130}$	
6	$-21.666^{+0.039}_{-0.040}$	$-3.946^{+0.027}_{-0.027}$	$-2.151^{+0.007}_{-0.007}$	—	37.711
	$-21.819^{+0.036}_{-0.034}$	$-4.224^{+0.023}_{-0.024}$	$-2.226^{+0.006}_{-0.006}$	$-5.232^{+0.097}_{-0.050}$	
7	$-22.226^{+0.088}_{-0.092}$	$-4.859^{+0.067}_{-0.070}$	$-2.483^{+0.012}_{-0.012}$	—	48.069
	$-22.104^{+0.076}_{-0.061}$	$-4.934^{+0.055}_{-0.046}$	$-2.522^{+0.011}_{-0.010}$	$-5.235^{+0.125}_{-0.048}$	
8	$-22.082^{+0.157}_{-0.135}$	$-5.307^{+0.134}_{-0.115}$	$-2.732^{+0.019}_{-0.016}$	—	11.431
	$-21.841^{+0.102}_{-0.101}$	$-5.281^{+0.085}_{-0.086}$	$-2.771^{+0.016}_{-0.015}$	$-5.179^{+0.216}_{-0.092}$	
9	$-21.224^{+0.174}_{-0.157}$	$-4.838^{+0.147}_{-0.135}$	$-2.702^{+0.024}_{-0.022}$	—	67.436
	$-19.023^{+0.018}_{-0.044}$	$-3.148^{+0.016}_{-0.036}$	$-2.304^{+0.029}_{-0.033}$	$-3.730^{+0.072}_{-0.080}$	
10	$-20.453^{+0.284}_{-0.242}$	$-4.768^{+0.320}_{-0.255}$	$-3.136^{+0.077}_{-0.046}$	—	3.410
	$-19.491^{+0.359}_{-0.632}$	$-3.900^{+0.362}_{-0.663}$	$-3.025^{+0.127}_{-0.142}$	$-4.136^{+0.193}_{-0.239}$	

Table 4.2: Best-fitting Schechter (first row corresponding to the redshift) and double power-law (second row corresponding to the redshift) function parameter values for the observed UV LF. The quoted error bars show the 16th – 84th percentile uncertainty obtained from the fit posteriors. We also provide the difference of the Bayesian Information Criterion (ΔBIC) value of the best-fitting parameters of the double power-law from the Schechter function.

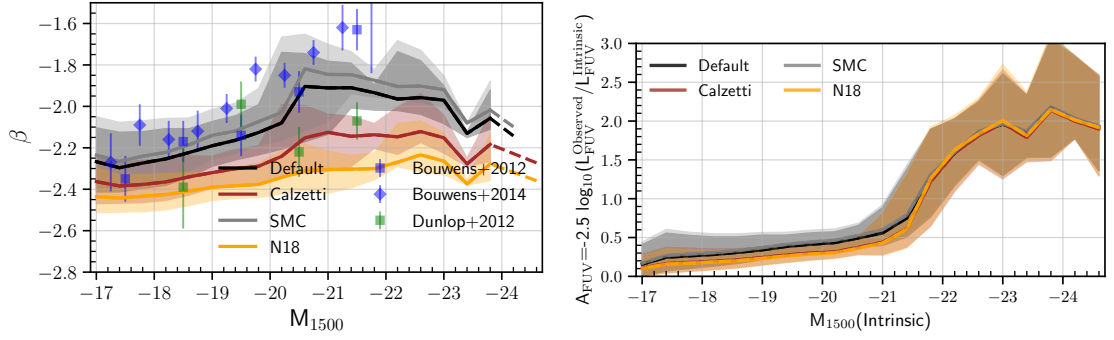


Figure 4.21: Left: Same as Figure 4.18, now showing β values for different extinction curves. Right: Attenuation in far-UV for different extinction curves at $z = 5$. Solid lines denotes the weighted median of the sample.

where the subscript i represents the bin of the property being measured, $n_{i,obs}$ is the number density of galaxies using the composite number density, $n_{i,exp}$ is the expected number density from the functional form being used (Schechter or double power-law), and σ_i is the error estimate. Using this form, σ can be explicitly provided by the expression, $\sigma_i = n_{i,obs} / \sqrt{N_{i,obs}}$, where $N_{i,obs}$ is the number counts in bin i from the re-simulations. We use flat uniform priors for the parameters in the functional forms. In the case of the double power-law form, to constrain the parameters, β was restricted to a lower limit of -5.3 and M^* to an upper limit of -19 .

For determining which functional form is better suited at different redshifts we calculated the Bayesian Information Criterion (BIC) value for the best-fit parameters. BIC is a criterion for model selection among a finite set of models, defined as follows:

$$BIC = -2 \ln(\mathcal{L}) + k \ln(N), \quad (4.11)$$

where \mathcal{L} is the likelihood of the fit function as expressed in Equation 4.10, k is the number of free parameters, and N is the number of data points used in the fitting. When fitting data, it is possible to increase the likelihood by adding more parameters, but can lead to overfitting. BIC resolves this by implementing a penalty term for the number of parameters in the model; the model with a lower BIC is preferred. A difference of ≥ 20 in the BIC value is usually taken to be a very strong preference for the model with a lower values. The difference of the BIC values, ΔBIC of the double power-law from the Schechter functional form is shown in Table 4.2.

4.C Other extinction curves

There has not been any consensus across observational or theoretical studies on the exact nature of the extinction curve in galaxies, since it is closely tied to the properties of the dust grains in galaxies. And this can be inferred better by probing the galaxy SED, and studies have suggested

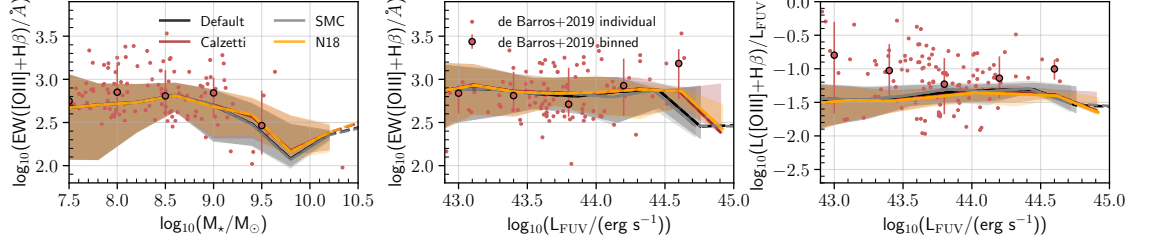


Figure 4.22: Same as Figure 4.13, now showing the line luminosity and equivalent widths for different extinction curves.

that using a single extinction curve for every galaxy might not be right. In our study we implement a simple extinction curve that is inversely proportional to the wavelength. In this section we will explore how some of the observables presented before changes depending on the chosen extinction curve, namely the Calzetti (Calzetti et al., 2000), Small Magellanic Cloud (SMC, Pei, 1992) and the curve used in (Narayanan et al., 2018, N18 from now on).

For this analysis we keep the value of κ_{BC} from our default model curve, i. e. $\kappa_{\text{BC}} = 1.0$. We then use the method described in Appendix 4.A to get κ_{ISM} , obtaining the values of 0.175, 0.0691 and 0.22 for the Calzetti, SMC and N18 curves respectively.

In the left panel of Figure 4.21 we present the effect of using different attenuation curves on the UV-continuum slope, β . It can be seen that the SMC curve has a higher median for β , compared to the default model, a consequence of the SMC curve being steeper than our default value. While for the case of the Calzetti and N18 curves, the former has a higher normalisation compared to the latter. We also tried increasing the value of κ_{BC} for the Calzetti and N18 curves to steepen the relation. We find that the match to the steepness of the observations is difficult to obtain from these curves, implying the FLARES galaxies prefer a steeper extinction curve similar to the SMC to reproduce the UV continuum observations.

In the right panel of Figure 4.21 we present the effect of using different attenuation curves on the attenuation in the far-UV. There is no observed difference in the attenuation in the FUV for any of the curves except at intrinsic $M_{1500} \gtrsim -21.5$ where the Calzetti and N18 curves produce on average lower attenuation. From our discussion before it is quite clear that despite this the underlying properties vary differently on using these different extinction curves.

In Figure 4.22 we present the effect of using different attenuation curves on the line luminosity and equivalent width relationship of the [OIII] λ 4959,5007 doublet. As can be seen all the curves trace the same space in all the sub-figures. Any minute difference seen happens at higher stellar mass/far-UV luminosity, with the default and SMC curve tracing a slightly lower median than the others.

FLARES III: THE PROPERTIES OF MASSIVE DUSTY GALAXIES AT COSMIC DAWN

5

Using the First Light And Reionisation Epoch Simulations (FLARES) we explore the dust driven properties of massive high-redshift galaxies at $z \in [5, 10]$. By post-processing the galaxy sample using the radiative transfer code `SKIRT` we obtain the full spectral energy distribution. We explore the resultant luminosity functions, IRX- β relations as well as the luminosity-weighted dust temperatures in the Epoch of Reionisation (EoR). We find that most of our results are in agreement with the current set of observations, but under-predict the number densities of bright IR galaxies (at $z = 5$), which are extremely biased towards the most overdense regions. We see that the FLARES IRX- β relation (for $5 \leq z \leq 8$) predominantly follows the local starburst relation. The IRX shows an increase with stellar mass, plateauing at the high-mass end ($\sim 10^{10} M_{\odot}$) and shows no evolution in the median normalisation with redshift. We also look at the dependence of the peak dust temperature (T_{peak}) on various galaxy properties including the stellar mass, IR luminosity and sSFR, finding the correlation to be strongest with sSFR. The luminosity-weighted dust temperatures increase towards higher redshifts, with the slope of the T_{peak} - redshift relation showing a higher slope than the lower redshift relations obtained from previous observational and theoretical works. The results from FLARES, which is able to provide a better statistical sample of high-redshift galaxies compared to other simulations, provides a distinct vantage point for the high-redshift Universe.

5.1 Introduction

The *Hubble Space Telescope* (*HST*) has been instrumental in the last decade observing the rest-frame UV of high-redshift galaxies (e. g. Beckwith et al., 2006; Bouwens et al., 2006; Wilkins et al., 2010; Robertson et al., 2010; Bunker et al., 2010; Wilkins et al., 2011b; Bouwens et al., 2014; Finkelstein et al., 2015; Bowler et al., 2017), finding more than 1000 galaxies at $z > 5$. These efforts from *HST* have been complemented by wide-area ground based near-IR surveys (e. g. *UltraVISTA*, Bowler et al., 2014; Stefanon et al., 2019) providing samples of rare bright galaxies. *Spitzer Space Telescope* observations (e. g. Ashby et al., 2013; Roberts-Borsani et al., 2016), probing the rest-frame optical at $z > 5$ has provided further constraints on these high-redshift systems. However, the UV/optical alone cannot unravel the nature as well as dynamical properties of these high-redshift systems, such as reliable estimates of the total star formation rates, since it is not an unbiased tracer due to the presence of dust.

Dust plays a major role in the observation of galaxies, with almost 30% of all photons in the Universe reprocessed by dust grains during their lifetime (Bernstein et al., 2002). Even though the average dust content of galaxies in the EoR is very low compared to the local Universe, it still has a significant impact on shaping observations by attenuating the emitted source radiation (Salim & Narayanan, 2020), particularly on the most massive galaxies. Thus it is crucial that we understand more about the effects of dust and how it affects the various observationally derived quantities. The stellar emission in a galaxy, which is predominantly in the UV-to-NIR, gets re-processed by the intervening dust into the IR regime. Over the years, the observations in this regime using far infrared (FIR), millimetre (mm) and sub-millimetre (sub-mm) observatories have been instrumental in mapping the dust content of galaxies. This has been done with the help of instruments like *ALMA* (e. g. Knudsen et al., 2017; Hashimoto et al., 2018; Smit et al., 2018; Bouwens et al., 2020), *Herschel* (e. g. Gruppioni et al., 2013; Wang et al., 2019), etc. In many cases there have been detections from deep *ALMA* and *PdBI* observations of galaxies at extremely high redshifts ($z > 6$) with large reservoirs of dust ($> 10^8 M_{\odot}$; Mortlock et al., 2011; Venemans et al., 2012; da Cunha et al., 2015).

The early identification of these dusty star-forming galaxies were from single-dish sub-mm surveys finding massive populations at $z > 1$ (see Casey et al., 2014). Even though they are rare, these galaxies contribute significantly to the cosmic star formation density during cosmic noon ($z \sim 2 - 3$, e. g. Bouwens et al., 2020; Zavala et al., 2021). The picture at higher redshift ($z > 4$) is still unclear. *ALMA* and *Herschel* have been instrumental in filling this space at high-redshift. Recent survey programmes like ALPINE (Le Fèvre et al., 2020; Béthermin et al., 2020; Faisst et al.,

2020a), ASPECs (Walter et al., 2016; Decarli et al., 2019; González-López et al., 2019), MORA (e. g. Zavala et al., 2021), etc are helping us to understand the dusty nature of high-redshift galaxies (also see Hodge & da Cunha, 2020, for more high- z surveys) by building a large statistical sample. High-redshift studies like Gruppioni et al. (2013); Wang et al. (2019); Gruppioni et al. (2020) have constructed IR luminosity functions (IR LF). The jury is still out on the normalisation of the IR LF at high-redshift due to difficulties in de-blending of IR data and smaller volumes probed in some surveys. Other observational studies like Schreiber et al. (2018); Bouwens et al. (2020) have explored the evolution of the luminosity-weighted dust temperatures, and have found an increase in the value with increasing redshift (up to $z \leq 5$). Another important observational space that has been studied widely is the relationship between the Infrared Excess (IRX) - UV continuum slope (β). Empirical relationships (e. g. Pettini et al., 1998; Meurer et al., 1999; Takeuchi et al., 2012; Reddy et al., 2015) built in this space using observations of low-redshift ($z \lesssim 2$) galaxies have been used to correct for dust attenuation in galaxies. There has been a variety of observational studies exploring this space at these high redshifts (e. g. Koprowski et al., 2018; Fudamoto et al., 2020; Bouwens et al., 2020; Schouws et al., 2021). They have found varying results that favours empirical relation using Calzetti as well as SMC extinction curve. The obtained relation is also strongly influenced by the adopted SED dust temperature and the functional form used to obtain the IR luminosity. With upcoming surveys on facilities like *JWST*, *Euclid*, *Roman Space Telescope*, and the *Atacama Large Aperture Submillimeter Telescope* (AtLAST, Klaassen et al. 2019) are expected to substantially contribute to these efforts to build a comprehensive picture of galaxy formation and evolution in the high-redshift Universe.

In addition to these observational efforts, it is crucial to study the nature of these dusty high-redshift systems using theoretical models of galaxy formation and evolution. Several studies have used techniques built with semi-analytical or analytical (e. g. Lacey et al., 2016; Popping et al., 2017b; Lagache et al., 2018; Lagos et al., 2019; Sommovigo et al., 2020) and hydrodynamical (e. g. Olsen et al., 2017; Narayanan et al., 2018; Ma et al., 2019; McAlpine et al., 2019; Liang et al., 2019; Baes et al., 2020; Trčka et al., 2020; Lovell et al., 2021b; Liang et al., 2021; Shen et al., 2021) models to explore and understand the trends and variations in observed properties of galaxies like the submillimeter number counts, infrared luminosity functions, IRX- β relations, dust temperatures, fine-structure transitions, etc. Many of them have been successful in reproducing various observational results, and has also been instrumental in understanding the underlying scaling relations.

In comparison to SAMs, hydrodynamical simulations model in greater detail the evolution of dark matter, gas, stars and black holes, allowing for a more detailed exploration of galaxy

structure and observed properties. A drawback of some of the current state-of-the-art cosmological hydrodynamic simulation periodic boxes is that they have fewer massive galaxies in the EoR, which are thought to be biased towards the most overdense regions (see Chiang et al., 2013; Lovell et al., 2018; Ito et al., 2020). This is mainly due to the unfeasible amount of computational time required to run much larger periodic volumes with the needed resolution to resolve the relevant scales at these high-redshifts. Hence they lack the statistical power to investigate the bright galaxies that will be discovered and investigated with the current or future generation of telescopes.

To overcome this dearth of a representative statistical sample of massive galaxies when studying the EoR, we use the First Light And Reionisation Epoch Simulations, FLARES; introduced in Lovell et al. (2021a); Vijayan et al. (2021) to study the dust driven properties of massive galaxies in the EoR. Using the radiative transfer code SKIRT (Camps & Baes, 2015) we post-process the galaxies to produce their full spectral energy distributions (SEDs). We aim to understand how well the EAGLE physics model is able to reproduce the high-redshift Universe, mostly in comparison to observations in the infrared part of the spectrum like the IR LF, IRX- β and luminosity-weighted dust temperatures. This work complements other theoretical studies in the high-redshift Universe and provide insights into how the intrinsic galaxy properties are connected to their observed and derived properties.

This chapter is structured as follows, in section §5.2 we introduce our galaxy sample and the method for SED generation. In section §5.3 we show our results, including the UV and IR luminosity in §5.3.1, the IRX- β relation in §5.3.2, and the variation and evolution of dust temperatures in §5.3.3. We finally summarise our findings and present our conclusions in section §5.4.

5.2 Methods

The FLARE simulation strategy has already been explained in Chapter 3. Here we will detail the selection of our galaxy sample, their physical properties and our spectral energy distribution modelling technique.

5.2.1 Galaxy Identification and Selection

Galaxies in FLARES, similar to the standard EAGLE, are identified with the SUBFIND algorithm (Springel et al., 2001; Dolag et al., 2009), which runs on bound groups found from via the Friends-Of-Friends algorithm (FoF, Davis et al., 1985). The galaxy stellar masses are defined using star

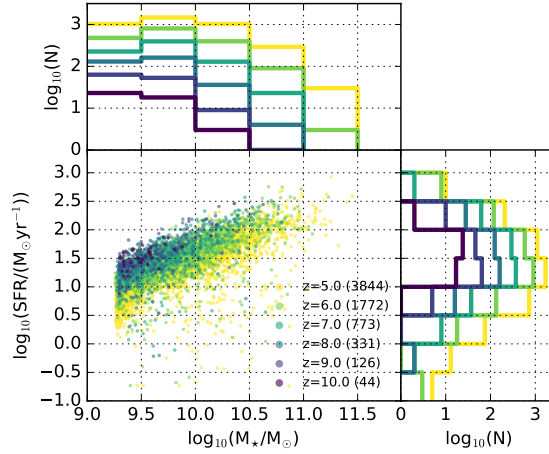


Figure 5.1: Shows the relationship between the galaxy stellar mass and the star formation rate (SFR) averaged over the star particles that were formed in the last 10 Myr for $z \in [5, 10]$. Also shown is the histogram of the distribution of stellar mass and SFR in different bins for these redshifts. The total number of galaxies at these redshifts are indicated within brackets alongside the legend.

particles within a 30 pkpc aperture centred on the most bound particle of the self-bound substructures. For the purpose of this study we concentrate only on the most well resolved galaxy systems that have more than 1000 star particles. This coincides with galaxies more massive than $\sim 10^9 M_\odot$ in stellar mass (see Figure 5.1). This selection also overlaps very well with the observationally inferred mass ranges (for e. g. the ALPINE survey; [Le Fèvre et al., 2020](#); [Béthermin et al., 2020](#); [Faisst et al., 2020a](#)) of the galaxies detected/followed up in the infrared with *ALMA* and other instruments.

In Figure 5.1 we plot the stellar mass of the selected galaxies against their star formation rate (SFR, quoted values are averaged for stars formed in the last 10 Myr) for $z \in [5, 10]$. We also show histograms of the galaxy stellar masses and SFR distributions. Our selection samples ~ 7000 galaxies in this redshift and mass regime. At $z = 10$, our sample of galaxies is only 44, and thus any inferences drawn can be subject to large scatter. The SFR seen in our selection has a maximum value just below $10^3 M_\odot/\text{yr}$.

5.2.2 Spectral Energy Distribution modelling

There are various methods to obtain the Spectral Energy Distribution (SED) of a galaxy in simulations. A comprehensive method to get them, so that the properties of the dusty medium is captured, is to perform radiative transfer. There are numerous codes (e. g. SUNRISE ([Jonsson, 2006](#)), RADMC-3D ([Dullemond et al., 2012](#)), SKIRT ([Camps & Baes, 2015](#)), POWDERDAY ([Narayanan et al., 2021](#)), etc) available, most relying on sophisticated Monte-Carlo methods. For this study we use the publicly available code SKIRT, version 9 ([Camps & Baes, 2020](#)).

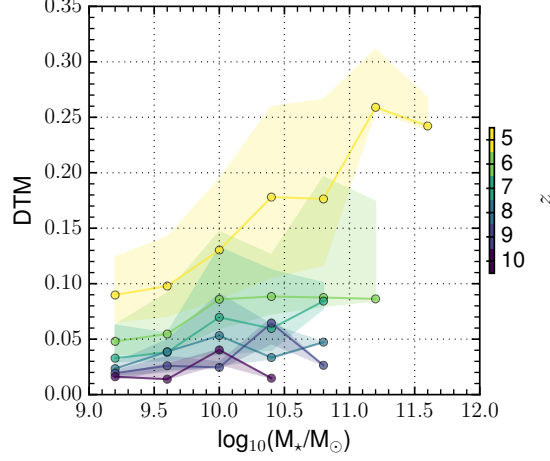


Figure 5.2: Shows the evolution of the dust-to-metal (DTM) ratio with the galaxy stellar mass across $z \in [5, 10]$. The solid line shows the weighted median while the shaded region indicate the 16-84 percentile spread in the value. The DTM decreases with increasing redshift.

FLARES does not inherently model dust formation and destruction, and thus cannot reliably estimate the amount, nature and distribution of dust in the different galaxies. For the purpose of obtaining the amount and distribution of dust we assume a constant dust-to-metal ratio ($\text{DTM} = M_{\text{dust}} / (M_{\text{metal}} + M_{\text{dust}})$) per galaxy, in SPH gas particles below temperatures of 10^6K or in star-forming gas particles. This temperature is higher than what was adopted in previous EAGLE-SKIRT work (e.g. [Camps et al., 2016](#); [Trayford et al., 2017](#)), and ensures that dust is only destroyed in the very hot gas phase in the galaxies. Changing the threshold to lower temperatures has negligible impact on the results presented in this work. The DTM ratio is calculated using the DTM fitting function in [Vijayan et al. \(2019\)](#), or Chapter 2, Equation 2.15), obtained from the dust model implemented in the L-Galaxies semi-analytical model. In that work, the DTM ratio is parameterised as a function of the mass-weighted stellar age and the gas-phase metallicity. Figure 5.2 shows the evolution and spread in the DTM ratio used in this work as a function of the galaxy stellar mass for $z \in [5, 10]$. It can be seen that there is an increase from a value of ~ 0.01 at $z = 10$ to ~ 0.2 by $z = 5$. The spread in the value also increases with decreasing redshift. More details on the evolution of the DTM ratio with redshift, and its dependence on various other galaxy properties, can be found in [Vijayan et al. \(2019\)](#). The use of a varying DTM ratio dependent on galaxy properties, as opposed to a constant value of 0.3, is another difference from previous EAGLE-SKIRT work. The evolution of the median DTM ratio with redshift seen here is similar to the one observed in [Vogelsberger et al. \(2020b\)](#) for ILLUSTRIS-TNG galaxies.

In this work we use the Small Magellanic Cloud (SMC) dust grain type and size distribution ([Weingartner & Draine, 2001](#)) built in to SKIRT. Due to its low-metallicity, the SMC is considered to be a good analogue to high-redshift galaxies. We use 8 grain size bins for silicate and graphite

dust types to compute the thermal emission. The dust grid for this setup is constructed using the built-in octree grid in SKIRT, using the previously defined dust particle distribution obtained from SPH gas particles. The octree is refined between a minimum refinement level of 6 and maximum of 16, with the cell splitting criterion set to a dust fraction value of 2×10^{-6} times the total dust mass in the domain, as well as a maximum V-band optical depth of 10. We use 10^6 photon packets per each radiation field wavelength grid, giving good convergence in observed properties. Our radiation field wavelength grid, as well as the dust emission grid, is spanned by a logarithmic grid between $0.08\text{--}1500\mu\text{m}$, with 200 points. We include dust self-absorption and re-emission in the set-up, with this procedure iterated such that the change in the absorbed dust luminosity is less than 3%. We place our detector to record the SED at a distance of 1Mpc enclosing a 60kpc a side square region. In this work we record multiple orientation sightlines, but the fiducial orientation is along the z-axis.

Similar to previous EAGLE-SKIRT work, we apply differing amount of dust attenuation to old and young stellar populations. Young stars, with stellar ages less than 10^7yr , are still embedded in their birth clouds, and as such experience higher dust attenuation (e. g. Charlot & Fall, 2000). We perform the same resampling technique that was employed in Camps et al. (2016); Trayford et al. (2017) to designate young and old stellar population from star particles and star-forming gas particles in the simulation. A difference from those works is that we do not subtract the contribution of dust from young stars which were part of the star-forming gas particles when we perform the resampling. This is because these particles already have gas/dust intrinsic to them (see section 2.4.4 in Camps et al., 2016, about introducing ‘ghost’ gas particles) unlike the resampled star particles which were converted to young stars. The emission from old stellar populations is modelled using the BPASS (Stanway & Eldridge, 2018) SPS library and the young stars with MAPPINGS III (Groves et al., 2008) templates. The former uses the Chabrier (Chabrier, 2003) IMF while the latter uses Kroupa IMF (Kroupa, 2002). We do not expect this difference to have a big effect since they are very similar. The BPASS model is characterised by the age and metallicity of the stellar particle while the MAPPINGS III template uses the SFR, metallicity, the pressure of the ambient ISM, the compactness of the HII region ($\log_{10}(C)$), and the covering fraction of the associated photo-dissociation region (f_{pdr}) of the star-forming particles. We use the same prescription for deriving the SFR, pressure and $\log_{10}(C)$ of the star particle as in Camps et al. (2016). However, we set the PDR covering fraction, f_{PDR} to 0.2, higher than 0.1 which was used in Camps et al. (2016). Our adopted value is same as the fiducial value used in Groves et al. (2008); Jonsson et al. (2010). A higher f_{PDR} results in more of the stellar emission to be absorbed by the dust present within the birth-clouds, implying that more of the light is re-processed to

the IR. Thus a higher value of f_{PDR} implies a higher value of the IR luminosity, but the exact nature of the change in the SED of the galaxy (for e. g. change in the position of where the peak IR emission is) also depends on the value of $\log_{10}(C)$ (also see §5.4 in [Liang et al., 2019](#)). For more details about the SKIRT set-up we have used, we refer the interested reader to [Camps et al. \(2016\)](#); [Trayford et al. \(2017\)](#).

We use the local thermal equilibrium set-up in SKIRT which means that the dust grains are in local equilibrium with the radiation field. This condition (as opposed to being in non-thermal equilibrium) will pre-dominantly affect the fluxes in the rest-frame mid-IR, but have very negligible effect on our predictions in this work (for more details see Appendix 5.A, where we have run SKIRT with the non-thermal equilibrium setup). We also include dust heating from CMB radiation, which at high-redshifts (since, $T_{\text{CMB}}(z) = T_{\text{CMB}}(z = 0) \times (1 + z)$; also see [da Cunha et al., 2015](#)) can be non-negligible. We do not include the effect of AGN on the SEDs (SKIRT has the capability to model AGN emission, see [Stalevski 2012](#); [Stalevski et al. 2016](#)); we will show in Appendix 5.B how the predictions are affected when adding the AGN bolometric luminosity to the infrared emission (the effect is negligible and only seen at the bright IR luminosity end). In a future work we will explore in more detail the effect of AGN on the UV emission from FLARES galaxies.

We had previously modelled the UV to near-IR SED of the FLARES galaxies in [Vijayan et al. \(2021, Chapter 4\)](#) using a line-of-sight (LOS) dust extinction model. That work calibrated the dust attenuation based on matching to the UV luminosity function and the UV luminosity–UV-continuum slope relation at $z = 5$, as well as the $[\text{OIII}]\lambda 4959, 5007 + \text{H}\beta$ equivalent width relation at $z = 8$. Here we do not perform any calibration, and only adopt the dust-to-metal ratio from the L-Galaxies SAM which was successful in reproducing many of the seen observational trends. This will enable us to better understand many of the successes and shortcomings of the EAGLE model when applied at high-redshift. We compare the UV luminosity of the galaxies from this model to the LOS model in Appendix 5.E.

5.3 Results

In this section we will look at what we can learn about the dust properties of massive high-redshift galaxies from the FLARE simulations, focussing on $z \in [5, 10]$. In §5.3.1 we will look at the infrared (IR) luminosity function, while exploring the $\text{IRX-}\beta$ space in §5.3.2. In §5.3.3 we will look at the dust temperatures of these galaxies, exploring both the SED–inferred as well as the peak dust temperatures. All these observables are also compared to current observations. It

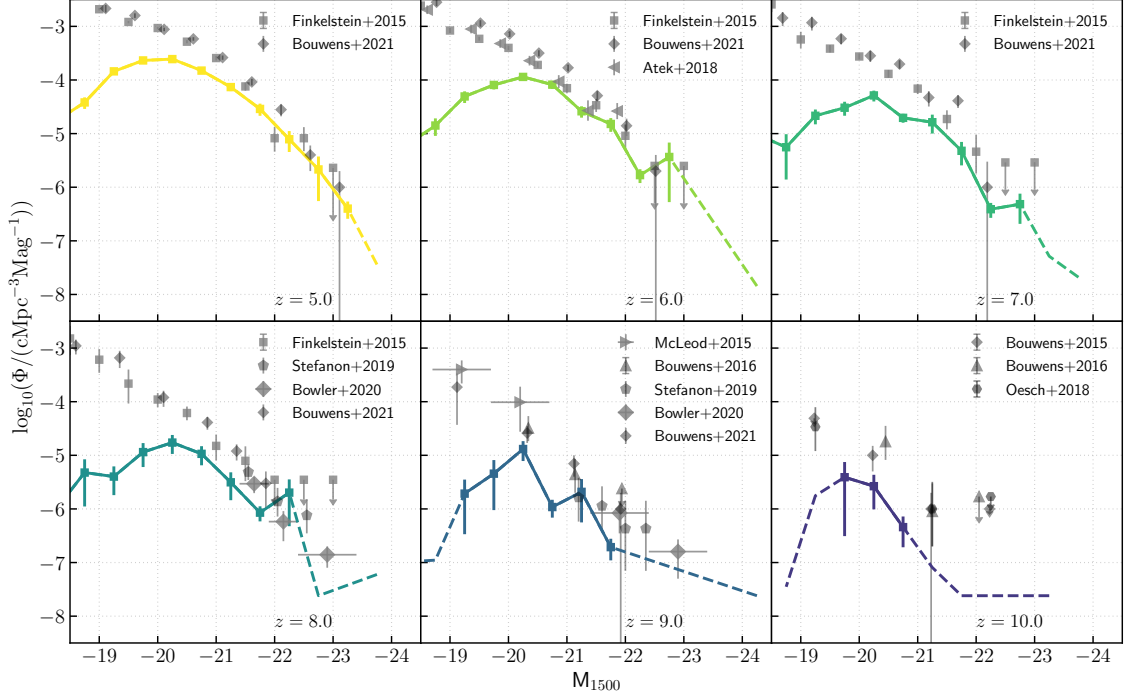


Figure 5.3: UV luminosity function of the FLARES galaxies for $z \in [5, 10]$. The errorbars show the Poisson $1-\sigma$ uncertainties for the different bins. Bins with fewer than 5 galaxies are represented by dashed lines. The data is incomplete at the faint-end due to our galaxy selection. We also plot alongside observational data from [McLeod et al. \(2015\)](#); [Finkelstein et al. \(2015\)](#); [Bouwens et al. \(2016, 2017\)](#); [Oesch et al. \(2018\)](#); [Atek et al. \(2018\)](#); [Stefanon et al. \(2019\)](#); [Bowler et al. \(2020\)](#); [Bouwens et al. \(2021\)](#).

should be noted that we do not model any observational effects (such as modelling the PSF or associated noise) that are inherent to the observed datasets that we compare to; this could impact derived properties and the associated systematic errors.

5.3.1 Luminosity functions

In Figure 5.3 we show the observed UV luminosity (measured at 1500\AA , plotted in magnitudes) function for the FLARES galaxies for $z \in [5, 10]$. This is an observational space where there is plenty of data and we compare our results to data from [McLeod et al. \(2015\)](#); [Finkelstein et al. \(2015\)](#); [Bouwens et al. \(2016, 2017\)](#); [Oesch et al. \(2018\)](#); [Atek et al. \(2018\)](#); [Stefanon et al. \(2019\)](#); [Bowler et al. \(2020\)](#); [Bouwens et al. \(2021\)](#). The UV LF is also usually used in calibration of dust models in high-redshift theoretical studies (e.g. [Wilkins et al., 2017](#); [Vogelsberger et al., 2020b](#); [Vijayan et al., 2021](#)). As can be seen, our model reproduces the UV LF reasonably well within the scatter seen in the observational data for $M_{1500} \lesssim -21$. The turnover at the faint-end is mainly due to our selection of well-resolved massive galaxies, whose contribution are at the bright end. It should be noted that there is a hint of galaxy number densities being slightly lower at $z = 8$ compared to the observations. There is also a similar trend at $z = 10$, but is harder to draw

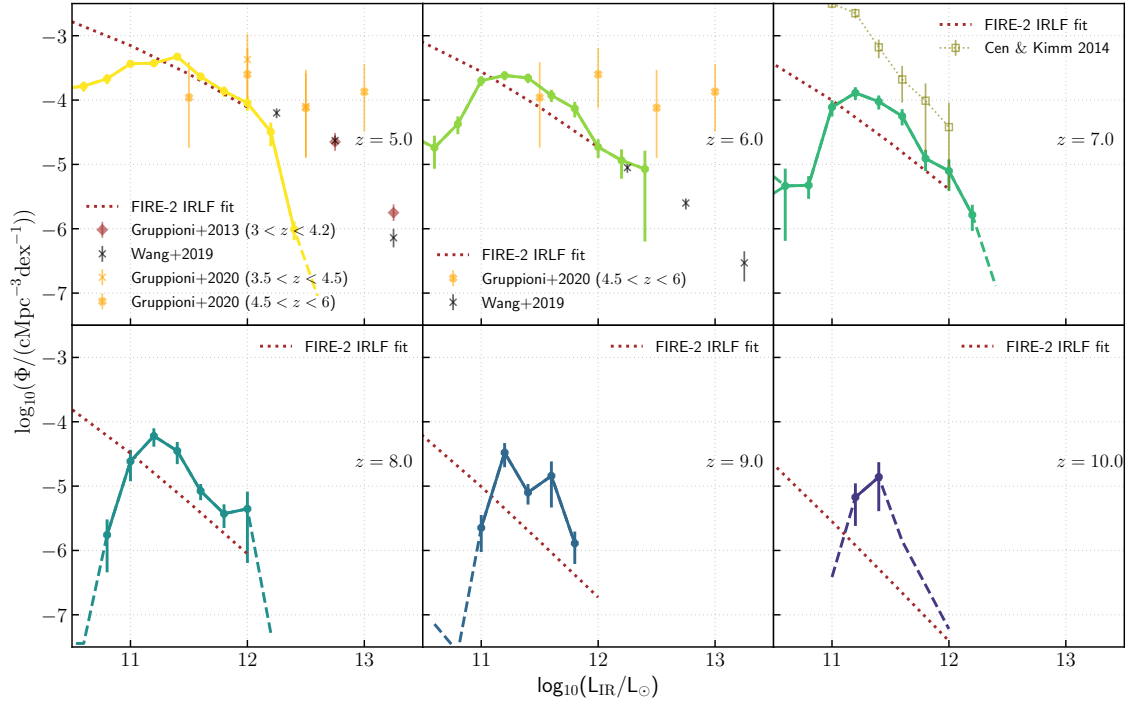


Figure 5.4: IR luminosity function of the FLARES galaxies for $z \in [5, 10]$. The errorbars show the Poisson $1\text{-}\sigma$ uncertainties for the different bins. Bins with fewer than 5 galaxies are represented by dashed lines. The data is incomplete at the faint-end due to our galaxy selection. We plot alongside observational data from Gruppioni et al. (2013); Wang et al. (2019); Gruppioni et al. (2020). Also plotted is the IR LF fit from the FIRE-2 simulations (Ma et al., 2019) and the $z = 7$ IR LF from Cen & Kimm (2014).

conclusions from, since the Oesch et al. (2018) data contains only 5 galaxies. Some of this tension can be attributed to the slightly lower normalisation (~ 0.3 dex) of the SFR function of the EAGLE reference volume or FLARES at intermediate SFR ($1 < \text{SFR} < 10 \text{ M}_{\odot} \text{yr}^{-1}$) as noted for high-redshift galaxies in Katsianis et al. (2017); Lovell et al. (2021a, also see Furlong et al. 2015) when compared to observed values. Vijayan et al. (2021) also showed that the unobscured SFR density of FLARES galaxies at $z \in [5, 7]$ showed slightly lower normalisation (~ 0.2 dex) in comparison with the unobscured value from Bouwens et al. (2020), even though the dust model was explicitly calibrated to match the UV LF, indicating that either the star formation rates are generally lower in the simulation, or the chemical enrichment rate (and thus the derived dust content) is higher, giving rise to higher attenuation than expected in these model galaxies. In the future with *JWST* we will be able to put tighter constraints on galaxy metallicities in the high-redshift regime. There is really good agreement at the high UV luminosity end at all the redshifts. Since the UV LF is predicted considerably well against observations (with the caveats noted) we will now try to draw meaningful conclusions from comparing against other observational spaces.

In Figure 5.4 we show the IR luminosity functions for $z \in [5, 10]$. The IR luminosity of the FLARES galaxies are obtained by integrating the observed SED between rest-frame wavelength

of $8 - 1000 \mu\text{m}$. We also plot alongside observational data from Gruppioni et al. (2013, using *Herschel* data); Wang et al. (2019, using the *Herschel* catalogue generated by the Bayesian source extraction tool XID+ in the COSMOS field); Gruppioni et al. (2020, using the the ALPINE-ALMA data) as well as theoretical results from FIRE-2 (Ma et al., 2019, their IR LF fit obtained from running SKIRT) and Cen & Kimm (2014, zoom simulation galaxy sample at $z = 7$, post-processed using SUNRISE (Jonsson, 2006)) for similar redshifts. It can be seen that FLARES is in agreement with the observational data for luminosities $\lesssim 10^{12} L_{\odot}$ for $z = 5$. There is a sharp decline in extremely bright IR galaxies in our simulation at $z = 5$. The very bright end of the function is under-estimated by ~ 1 dex compared to Gruppioni et al. (2013, 2020), which are collated measurements within broad redshift ranges. Due to this broader redshift range, the normalisation can be higher, since lower redshifts are expected to have higher number densities. However, a difference of ~ 1 dex is in tension with our predictions. Zavala et al. (2021) have also described the IR LF measurements in Gruppioni et al. (2020) to be representative of an overdense patch in the high-redshift Universe. This inference comes from the observational targets being highly clustered massive galaxies ($\log_{10}(M/M_{\odot}) \gtrsim 10.5$). This is in good agreement of the plotted IR LF of the overdense regions shown in Figure 5.6. In case of the Wang et al. (2019) data at $z = 5$ there is a similar case of underprediction of bright IR luminous galaxies. Thus we are inconsistent in a regime where two independent measurements (Gruppioni et al., 2013; Wang et al., 2019) agree, though it should also be noted that they are both obtained from the *Herschel* catalogue, and are subject to uncertainties associated with the deblending techniques employed. Thus they can be ideally treated as upper-limits on the IR luminosity function.

A reason for this sudden decrease is that our extreme IR-bright galaxies are biased towards the most overdense regions, having much lower contribution to the IR LF (see §5.3.1.1). Following from our argument in the UV LF section, the lower normalisation of the star formation rate function in our model at these redshifts also contributes to this lower number density (also discussed in McAlpine et al., 2019; Baes et al., 2020). This has also been investigated at lower redshifts and has been similarly attributed to the lower star formation rate as well as the lack of ‘bursty’ star formation in the EAGLE model (see McAlpine et al., 2019). However, our result is not an isolated case and has been a feature of many other cosmological and zoom simulations like ILLUSTRIS-TNG (Shen et al., 2021) and FIRE-2 (Ma et al., 2019, also plotted in Figure 5.4) at these redshifts. The SIMBA (Davé et al., 2019) suite of simulations shows a higher normalisation of the SFR function than EAGLE at high-redshift. In Lovell et al. (2021b) (where they post-processed SIMBA galaxies using POWDERDAY, (Narayanan et al., 2021)), they find reasonable agreement with observationally inferred $850\mu\text{m}$ number counts, which they partly attribute to the higher SFRs.

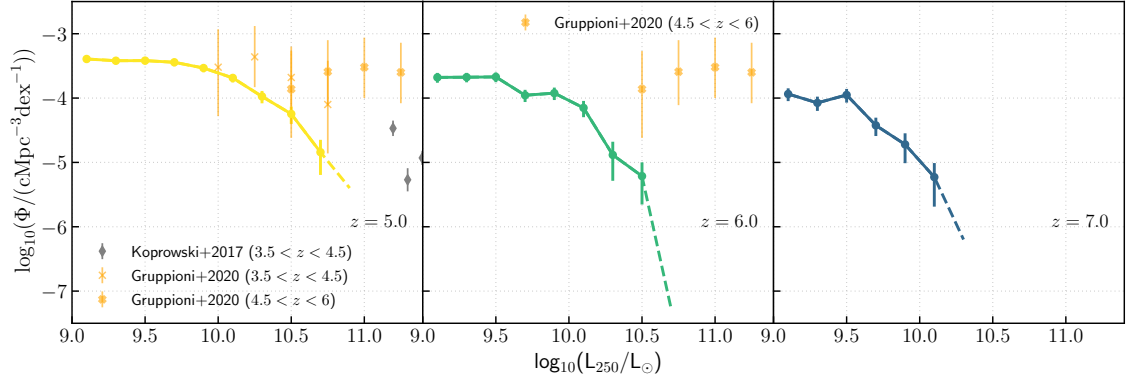


Figure 5.5: The rest-frame 250 μ m luminosity function of the FLARES galaxies for $z \in [5, 7]$. The errorbars show the Poisson 1- σ uncertainties for the different bins. Bins with fewer than 5 galaxies are represented by dashed lines. The data is incomplete at the faint-end due to our galaxy selection. We plot alongside observational data from [Koprowski et al. \(2017\)](#); [Gruppioni et al. \(2020\)](#) at similar redshift range.

So the lower star formation rate at high-redshift is a likely cause of the deficit in IR luminosities in our model.

There are caveats that come along with physics recipes to produce higher star formation. For example, there is a lack of quiescent galaxies at high-redshifts in SIMBA compared to observations and EAGLE, as explored in [Merlin et al. \(2019\)](#). A fine interplay of feedback and star formation is fundamental to match the various observational results and thus provide test beds for improving the model recipes. There has also been suggestions of changes to the initial mass function to a top heavy one in the most luminous galaxies to produce the seen higher number density of IR luminous galaxies (see for e.g. [Motte et al., 2018](#); [Schneider et al., 2018](#); [Zhang et al., 2018](#)). Any of these two scenarios would imply higher dust content from increased star-formation, thus reconciling the increase in the intrinsic emission with higher attenuation.

On comparing the FLARES IR LF at $z = 6$, we are in very good agreement with the [Wang et al. \(2019\)](#), also similar to what was seen for the IR LF of EAGLE galaxies in that study) and FIRE-2 results, while still being more than ~ 0.5 dex lower compared to the [Gruppioni et al. \(2020\)](#) data for $4.5 \leq z \leq 6$. We are in agreement with the FIRE-2 IR LF in the overlapping region. At $z = 7$, we compare to the theoretical predictions from the radiative transfer calculations of galaxies done in [Cen & Kimm \(2014\)](#). They show a higher normalisation, with their relation having a steeper faint-end evolution similar to our results at the extreme bright end. The higher normalisation can be attributed to the zoom simulation region representing 1.8σ matter density fluctuation on the chosen volume.

In Figure 5.5 we plot the rest-frame 250 μ m luminosity function of the FLARES galaxies in $z \in [5, 7]$. We also compare to observational data from [Koprowski et al. \(2017\)](#), using sub-mm/mm

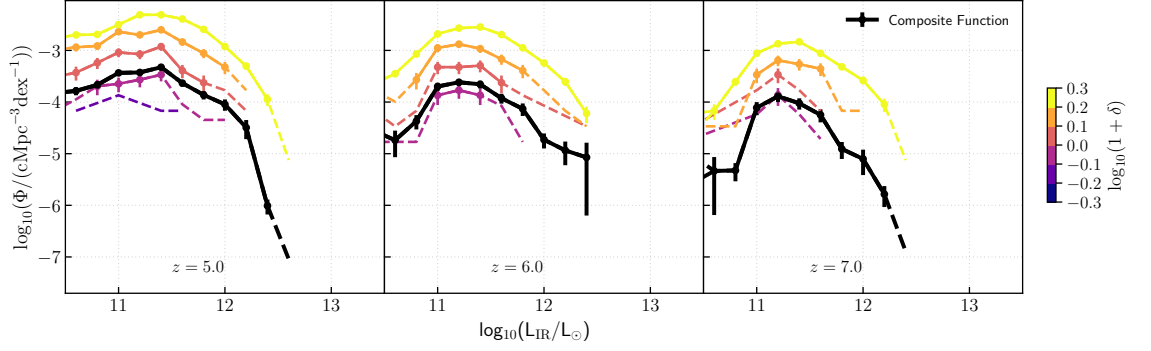


Figure 5.6: The FLARES IR LF for $z \in [5, 7]$ split by binned log-overdensity. Error bars denote the Poisson $1\text{-}\sigma$ uncertainties for each bin from the simulated number counts. The composite distribution function is plotted as black solid line.

imaging from SCUBA-2 and ALMA); Gruppioni et al. (2020). In general, we see an underprediction of the rest-frame $250\mu\text{m}$ luminosity function. We are in agreement with the $3.5 < z < 4.5$ Gruppioni et al. (2020) data within the uncertainties, while at $z = 6$, our results are ~ 1 dex lower at the extreme bright end. It can also be seen that the Gruppioni et al. (2020) data in the two redshift range show no clear decline in the number density galaxies, while in our case there is a reduction in number density by ~ 0.5 at the bright end of the function. But this is inconsistent with Koprowski et al. (2017) where they found a lower number density in the range $3.5 < z < 4.5$ compared to the Gruppioni et al. (2020) values. This discrepancy in the two data sets could be due to incompleteness in the sample selection associated with the Koprowski et al. (2017) data as well as the Gruppioni et al. (2020) data being representative of an overdense region. Nevertheless, our data does not extend to the extreme luminosities that the Koprowski et al. (2017) sample covers.

In Appendix 5.B we add the AGN bolometric luminosity to the IR luminosity for comparison to observations to gauge the effect AGN has on the IR LF. We see very small changes, not enough to reconcile an order of magnitude difference at the bright end of the IR LF. Also, in Appendix 5.C, we look at how our results are consistent with the current setup when excluding emission from birth clouds of young stars for $z \geq 8$.

5.3.1.1 Environmental Dependence of IR LF

In Figure 5.6, we show the IR LF in different matter overdensity bins for $z \in [5, 7]$. The composite function sits within the boundary of the positive and negative overdensity bins as expected. The plot shows that there is an increase in the number densities of IR luminous galaxies with increasing overdensity. The smallest overdensity bin is not visible in the plot since it is below the plotted IR luminosity range. At $z = 5$, there is an increase of ~ 1.5 dex in number density of

galaxies with IR luminosity of $\sim 10^{11}L_{\odot}$. This is very similar across rest of the plotted redshift range. At $z = 5$, only the most overdense regions contribute to the very bright end of the IR LF, which results in the rapid fall in the number densities seen in the composite function.

5.3.2 IRX- β

In this section we will look at the infrared excess (IRX) of the galaxies in the FLARE simulation. The infrared excess is defined as the ratio of the total infrared luminosity (L_{IR}) over the UV luminosity (L_{UV}), and can be derived as follows

$$\text{IRX} = \frac{L_{\text{IR}}}{L_{\text{UV}}} \simeq \frac{\int_{8\mu\text{m}}^{1000\mu\text{m}} L_{\lambda} d\lambda}{L'_{1500}}, \quad (5.1)$$

where $L'_{1500} = L_{1500} \times 1500\text{\AA}$, with L_{1500} being the far-UV luminosity calculated at 1500\AA . The UV-continuum slope, β is defined such that $f_{\lambda} \propto \lambda^{\beta}$ or alternatively $f_{\nu} \propto \lambda^{\beta+2}$, for λ in the rest-frame UV range. We measure β using the following prescription,

$$\beta = \frac{\log_{10}(L_{1500}/L_{2500})}{\log_{10}(1500/2500)} - 2, \quad (5.2)$$

where L_{1500} and L_{2500} are the far-UV and near-UV luminosity, respectively.

The IRX- β relation has been explored in numerous theoretical (e. g. [Safarzadeh et al., 2017](#); [Ma et al., 2019](#); [Trčka et al., 2020](#); [Schulz et al., 2020](#); [Liang et al., 2021](#)) and observational studies (e. g. [Reddy et al., 2015](#); [Wang et al., 2018](#); [Fudamoto et al., 2020](#); [Bouwens et al., 2020](#)) in low and high-redshift galaxies. Some studies have provided empirical relations for this plane assuming a dust screen model. These relations are expected to arise from the simple assumption that with increasing dust attenuation the UV-continuum slope becomes redder with the L_{IR} to L_{UV} ratio increasing. The assumption of different dust attenuation curves will determine the trajectory of this relation. However, in galaxies one would expect different attenuation for young and old stars, as well as different dust distribution across the galaxy that can provide large variation to the relationship (e. g. [Narayanan et al., 2018](#); [Schulz et al., 2020](#)).

The empirical relation between IRX and β is widely used to correct for the amount of dust-obscured star formation within galaxies at high-redshift. This is based on the assumption that high-redshift galaxies follow the same relation as their local analogues. However, some high-redshift galaxies seem to have smaller IRX than their local analogues (e. g. [Capak et al., 2015](#); [Fudamoto et al., 2020](#)). This has been largely attributed to the low dust mass temperatures adopted in modelling the observational data (e. g. [Sommovigo et al., 2020](#); [Bouwens et al., 2020](#)). Another cause of concern in using this relation at high redshift is the observed spatial offset

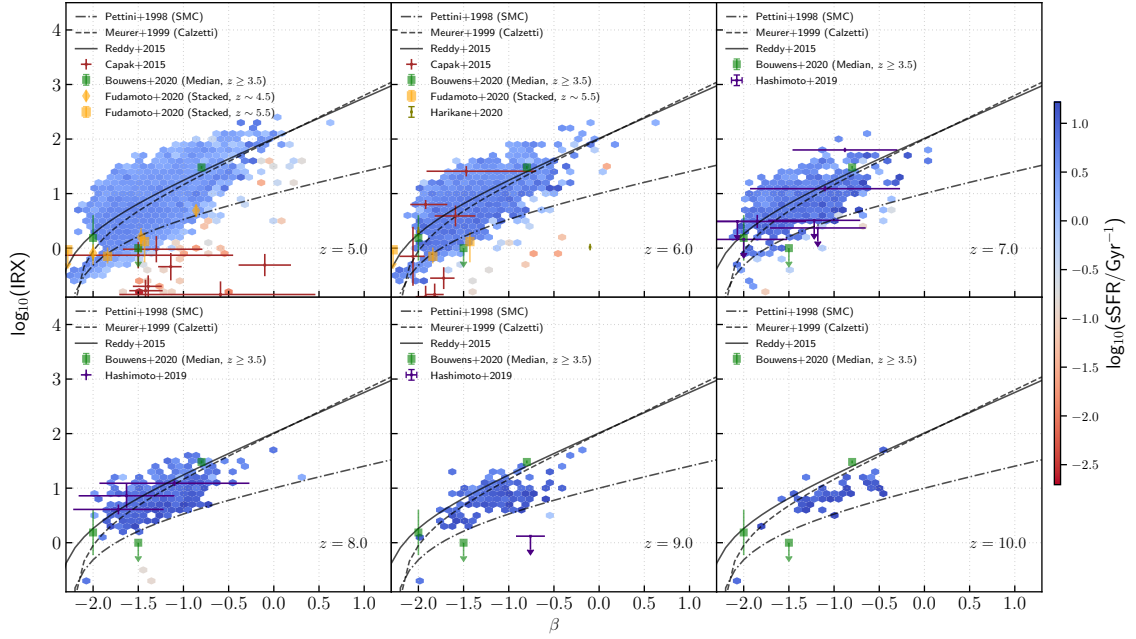


Figure 5.7: IRX- β distribution of the FLARES galaxies for $z \in [5, 10]$. The hexbins are coloured by the median specific star formation rate. We also show individual observational data at similar redshift from Capak et al. (2015, updated values from Barisic et al. 2017), Hashimoto et al. (2019, also with observations collated from Ouchi et al. 2013; Ota et al. 2014; Inoue et al. 2016; Knudsen et al. 2017; Laporte et al. 2017; Hashimoto et al. 2018; Marrone et al. 2018; Smit et al. 2018; Tamura et al. 2019). We also show stacked results from Fudamoto et al. (2020); Bouwens et al. (2020). Also shown is the empirical relation for dust screen models for SMC (Pettini et al., 1998) and Calzetti (Meurer et al., 1999), and from Reddy et al. (2015).

between the UV and IR emission (e. g. Bowler et al., 2018), possibly due to the birth cloud dispersal time in galaxies (Sommovigo et al., 2020).

We look at the variation of the IRX- β relation across $z \in [5, 10]$ in Figure 5.7. The plane is represented by hexbins which are coloured by their median sSFR values. Due to our simulations containing a large selection of extreme overdensities, there is an overabundance of IR luminous dusty galaxies in our data. We also plot observational data at similar redshifts from Capak et al. (2015, with updated values from Barisic et al. 2017, the IR luminosity was obtained using a power-law + MBB functional form using the following ranges: $\alpha = 1.5 - 2.5$ (mid-IR power-law slope), $\beta = 1.2 - 2.0$ (emissivity index) and $T_d = 25 - 45$ K (black body temperature), see equation 5.5), Hashimoto et al. (2019, also with observations collated from Ouchi et al. 2013; Ota et al. 2014; Inoue et al. 2016; Knudsen et al. 2017; Laporte et al. 2017; Hashimoto et al. 2018; Marrone et al. 2018; Smit et al. 2018; Tamura et al. 2019, all obtained using optically-thin MBB function with $T_d = 50$ K and $\beta = 1.5$, see equation 5.8), Harikane et al. (2020, obtained by fitting observed fluxes using optically-thin MBB with $\beta = 1.6$ and varying the IR luminosity and dust temperature) as well as the stacked and median results from Fudamoto et al. (2020, estimated using the conversion factor from $158 \mu\text{m}$ to L_{IR} presented in Béthermin et al. (2020)

using stacking of sources; caveats of the conversion being that it could be largely inaccurate for outliers with extreme dusty SEDs as well as stacking favouring brighter sources) and [Bouwens et al. \(2020\)](#), using optically-thin MBB with $\beta = 1.6$ and redshift evolution of dust temperature based on their equation 1). Also shown is the empirical IRX- β relations from [Pettini et al. \(1998, SMC\)](#), [Meurer et al. \(1999, Calzetti\)](#) and [Reddy et al. \(2015\)](#).

From the figure we can see that FLARES lies within the scatter of the observational values. There are a few exceptions in case of data with high- β and low-IRX values (low dust content and older stellar populations) found for a few galaxies in [Capak et al. \(2015\)](#); [Hashimoto et al. \(2019\)](#); [Harikane et al. \(2020\)](#). This could be a drawback of our implemented model that does not fully capture the diverse star-dust geometries of galaxies in these observations. These galaxies having high- β and low-IRX imply that they are moving away from the main-sequence relation towards quiescence. The FLARES sample could be inefficient in producing such galaxies. The quiescent galaxy population in FLARES will be probed in a future work where their number densities will also be explored. Also to be noted in case of the [Capak et al. \(2015\)](#) values, the upper limit for the dust SED temperature (in our equation 5.8) is ~ 20 K lower than our median values and hence the obtained IR luminosity will be lower. A similar dearth of high- β , low-IRX galaxies is seen in [Ma et al. \(2019\)](#). However, their probed galaxy stellar mass range is lower than ours and thus it could also be due to the fact that there are not any low-sSFR high-mass galaxies, that are moving into the quiescent regime.

At high-redshifts ($z > 4$), no clear picture has emerged on what kind of attenuation relation galaxies follow, whether or not they are consistent with the local relation, following a starburst or Calzetti-like attenuation law, or a shallower one like the SMC. As can be seen in Figure 5.7, the massive galaxies in FLARES predominantly follow the Calzetti or [Reddy et al. \(2015\)](#) like relation, with a small proportion of galaxies following the SMC relation at $z \in [5, 8]$, even though we adopted the SMC grain distribution. We also note that the galaxies which drop below the canonical relations are the ones that exhibit very low sSFR values, or with older stellar populations, in agreement with theoretical studies like [Narayanan et al. \(2018\)](#). At higher redshifts ($z \geq 8$) there is a hint of some transition away from the Calzetti relation towards the SMC one in FLARES, which we fail to properly capture due to the limited mass resolution of our simulation. This leads to the lack of well resolved low-mass galaxies in our sample to populate this space.

The reason for the majority of FLARES galaxies following the local starburst relation could be due to the inhomogeneous nature of dust at high redshift, with the β values being dominated by unobscured young stars while the IRX is dominated by dust emission near the highly obscured dust patches. With our selection of the most massive galaxies, this is the most likely outcome due

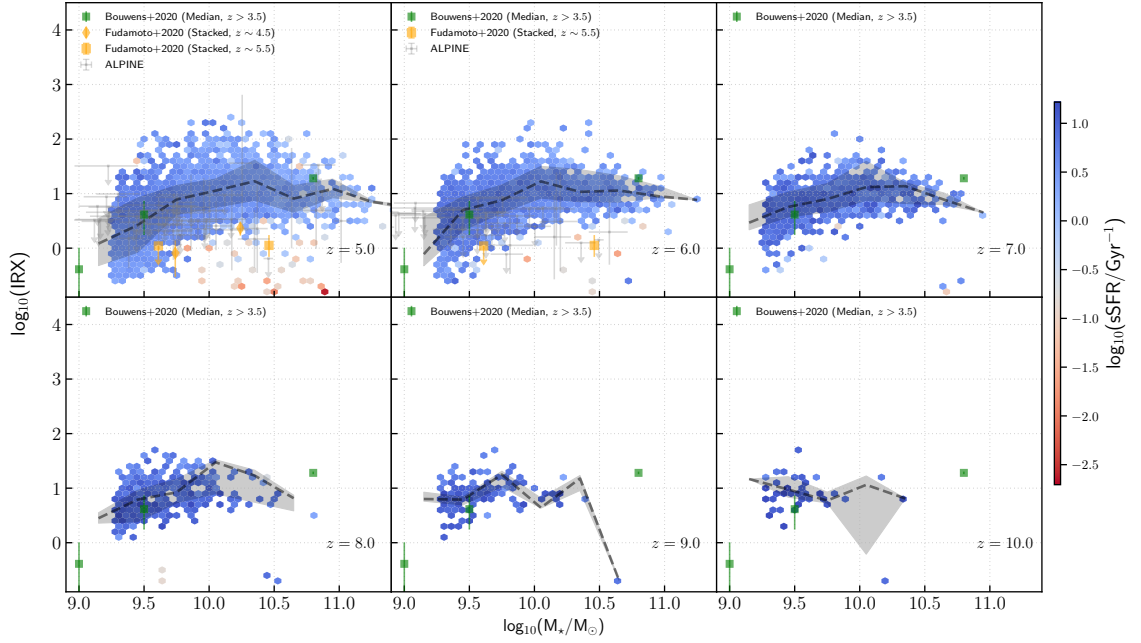


Figure 5.8: IRX-stellar mass distribution of the FLARES galaxies for $z \in [5, 10]$. The hexbins are coloured by the median specific star formation rate. The dashed line and the shaded region denote the weighted median and the 16-84 percentile spread of the data. We plot observational data from the publicly available ALPINE collaboration data (Le Fèvre et al., 2020; Béthermin et al., 2020; Faisst et al., 2020a). Also shown is the stacked and median relation from Fudamoto et al. (2020) and Bouwens et al. (2020) respectively.

to their higher dust content. Narayanan et al. (2018), using cosmological zoom simulations run using GIZMO (Hopkins, 2015), also attributed the major drivers of the observed deviations from the canonical relation to older stellar populations, complex star-dust geometries and variations in dust extinction curves. A similar result was also seen at high-redshift in Schulz et al. (2020), studying the IRX- β relation in ILLUSTRIS-TNG galaxies at $z = 0-4$. They concluded that the seen deviations could be best described in terms of sSFRs driving the shift in β with the star formation efficiency possibly being a good indicator of the variations in star-dust geometry. Liang et al. (2021) explored in detail the secondary dependencies of the IRX- β relation, concluding that the main driver of the scatter is the variations in the intrinsic UV spectral slope and thus the age of the underlying stellar population. Thus, due to large degeneracies among sSFRs or ages, star-dust geometry as well as dust compositions, it would be hard to pin-point a global track in the IRX- β relation for galaxies at these redshifts for different stellar masses.

In Figure 5.8 we look at the relationship between the galaxy stellar mass and IRX. We also plot observational results from the publicly available dataset of the ALPINE collaboration (Le Fèvre et al., 2020; Béthermin et al., 2020; Faisst et al., 2020a, values obtained from SED fitting) as well as the stacked and median results from Fudamoto et al. (2020) and Bouwens et al. (2020) respectively. Also shown is the weighted median and the 16-84th percentile variation for the

sample of galaxies. This plane provides further insights into what we saw in Figure 5.7. We can see that the galaxies following the SMC relation have stellar masses of $\sim 10^9 M_\odot$. These galaxies are the ones in the process of transitioning to the higher attenuation relation from rapid dust enrichment (see Figure 4.9 and 4.10 in Chapter 4 where we also see a rapid rise in the UV attenuation $\sim 10^9 M_\odot$ in stellar mass). We also see that there is a general trend towards high median IRX values with higher stellar masses, plateauing or slightly dropping at the highest masses. If we extrapolate the median to lower masses, it would lead to lower IRX values. This would indicate that the region near the SMC relation would be occupied by the lower mass galaxies. This is in agreement with what was seen in Ma et al. (2019) using the FIRE-2 simulation, where the majority of the galaxies below a stellar mass of $10^9 M_\odot$ follow the SMC relation.

The trend at the massive end ($\sim 10^{10} M_\odot$), which is clear at $5 \leq z \leq 8$, points towards an increase in the UV luminosity not being reflected to the same extent in the IR luminosity. This points towards decreasing dust attenuation in the most massive galaxies. This was also seen in the LOS dust attenuation model applied on the FLARES galaxies in Figure 4.10. The observed UV LF being better fit by double power-law at these high-redshift (e. g. Bowler et al., 2014, 2020; Shibuya et al., 2021) also points towards decreasing dust attenuation at the bright and massive end, that can contribute to this decline. However, studies such as Ferrara et al. (2017) posit that some of the IRX deficit galaxies could have dust embedded in large gas reservoirs, thus not contributing to an increase in the IR luminosity.

On comparing to the observational data, the median values from Bouwens et al. (2020, note that the highest mass bin has just one galaxy) are a good match. In case of the ALPINE data as well as the stacked results from Fudamoto et al. (2020, which also uses ALPINE data), our median relation is higher than their dataset. However, this can be explained by the UV selection of the galaxies observed in the survey, which can miss the dustier systems.

There is no noticeable evolution of IRX-stellar mass relation with redshift. The normalisation of the median value does not show any redshift evolution.

5.3.3 Dust temperatures

In this section we will look at the dust temperatures of the galaxies in FLARES. There are different definitions of the dust temperature, both observational and theoretical. Here we will look at the SED-derived and peak wavelength temperature, which are both measures of the light-weighted dust temperatures, but can be considerably different depending on the functional forms being used (see e.g. Casey, 2012). These measures are very different to the mass-weighted temperature,

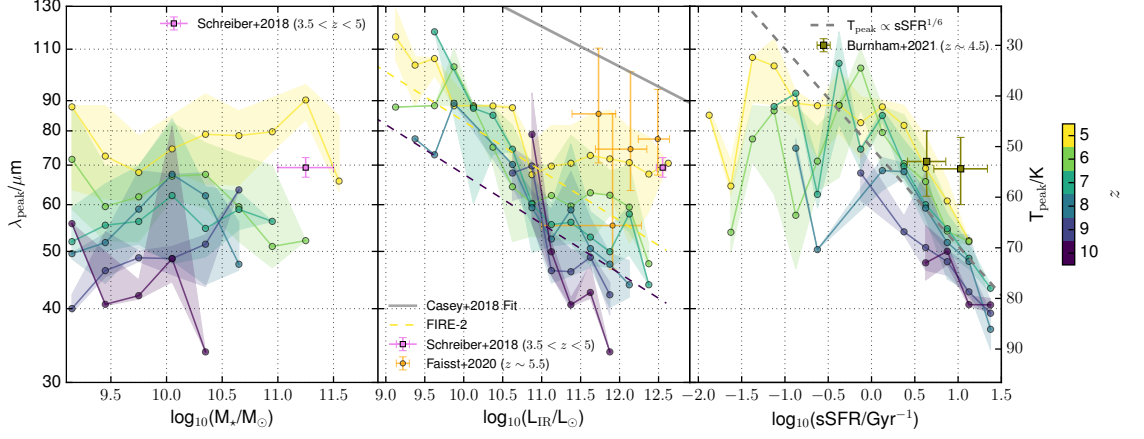


Figure 5.9: We show the variation of λ_{peak} (corresponding T_{peak} values are shown on the right y-axis) with various galaxy properties like the galaxy stellar mass (left panel), total infrared luminosity (middle panel) and the specific star formation rate (sSFR, right panel) for $z \in [5, 10]$. The weighted median (solid line) and the 16-84 percentile (shaded region) variation is plotted for the different redshifts which are denoted by the different colours as per the shown colourbar. We also show observational data from Schreiber et al. (2018, median data in $3.5 < z < 5$) for the stellar mass range of $10^{11} - 10^{11.5} M_{\odot}$ in the stellar mass- λ_{peak} plane and the measurements from Faisst et al. (2020b, for $z \sim 5.5$) in the IR luminosity- λ_{peak} plane. The fits obtained from observations of $z < 5$ galaxies from Casey et al. (2018, grey solid line) and FIRE-2 (Ma et al., 2019, for $z = 5, 10$, with same colour as the FLARES median lines, but dashed) simulations are also shown. We also plot a $T_{\text{peak}} \propto \text{sSFR}^{1/6}$ relation to compare to our median relations in the right panel.

a measure of mostly the cold dust content of galaxies, which is expected to be largely independent of redshift and galaxy properties as well as significantly lower than the light-weighted temperatures (Liang et al., 2019; Sommovigo et al., 2020).

5.3.3.1 T_{peak}

The peak dust temperature (T_{peak}) can be obtained from the Wein displacement law from the rest-frame wavelength at which the infrared flux density peaks (λ_{peak}). This is defined as

$$T_{\text{peak}} = \frac{2.898 \times 10^3}{\lambda_{\text{peak}}} \mu\text{m K}, \quad (5.3)$$

which follows the relation for a true blackbody (that has a dust emissivity index β of 2). This measure has been used in many observational studies to understand the evolution of luminosity weighted dust temperature across redshifts and galaxy properties (e. g. Casey et al., 2018; Schreiber et al., 2018; Burnham et al., 2021). It suffers less from model dependent biases compared to other dust temperature values. It should also be noted that T_{peak} is only a proxy for λ_{peak} , and the choice of the normalisation is to compare with other theoretical and observational studies.

In Figure 5.9, we show how the weighted median as well as the 16-84th percentile (shaded region) peak of the IR emission varies with various galaxy properties like the stellar mass (left

panel), IR luminosity (middle panel) and the specific SFR (sSFR, right panel) for $z \in [5, 10]$. In case of the variation in λ_{peak} with galaxy stellar mass, we do not find any significant trend in the ranges we are considering. There is a general decrease (increase) in the λ_{peak} (T_{peak}) value with increasing redshift. The observational data from [Schreiber et al. \(2018\)](#), which shows the median data in $3.5 < z < 5$ in the stellar mass range $10^{11} - 10^{11.5} M_{\odot}$, is in agreement with our data at $z = 5$.

In a similar vein, the middle panel of Figure 5.9 shows the variation of λ_{peak} with the IR luminosity for different redshifts. We also plot along with it observational data from [Schreiber et al. \(2018\)](#), same range as before) and [Faisst et al. \(2020b\)](#), for galaxies at $z \sim 5.5$) as well as the redshift independent relation presented in [Casey et al. \(2018\)](#) for $z < 5$ samples. Also shown is the redshift dependent fit to the FIRE-2 ([Ma et al., 2019](#), post-processed with SKIRT) for $z = 5, 10$. The data from [Strandet et al. \(2016\)](#); [Faisst et al. \(2020b\)](#) match well with our constraints at $z = 5$, while the relation from [Casey et al. \(2018\)](#) is well above our values as well as the high-redshift observations. The FLARES galaxies agree well with the $z = 5$ fit from FIRE-2, but the slope of the relation at higher redshifts ($z \geq 7$) is steeper in FLARES. There is a trend of lower (higher) values of λ_{peak} (T_{peak}) with increasing IR luminosity similar to the what is seen in [Ma et al. \(2019\)](#) and [Shen et al. \(2021\)](#), with the SKIRT, post-processed ILLUSTRIS-TNG galaxies). However, by $\sim 10^{11} L_{\odot}$ we see a flattening in this relation similar to what was found in [Shen et al. \(2021\)](#) at $z = 4, 6$, with a higher normalisation than the one here. As posited there as well as in other studies ([Jin et al., 2019](#)), this trend could be due to the increasing optical depth in the most luminous galaxies, hiding the warm dust associated with star-forming compact regions, making the contribution to the dust temperature minimal. At the high IR luminosity end, there is a strong evolution towards lower (higher) λ_{peak} (T_{peak}) values with redshift, showing that FLARES also prefers an evolving relation similar to results in [Ma et al. \(2019\)](#); [Shen et al. \(2021\)](#).

The right panel of Figure 5.9 shows the variation of λ_{peak} with the sSFR (SFR calculated using stars born in the last 10 Myr) for different redshifts. We have also over-plotted two galaxies at $z \sim 4.5$ from [Burnham et al. \(2021\)](#) which are in agreement with our $z = 5$ relation within the scatter. We see a very tight relation for the FLARES galaxies at the high-sSFR ($\text{sSFR}/\text{Gyr}^{-1} \gtrsim 0$) end. This has been observed in studies like [Magnelli et al. \(2014\)](#); [Ma et al. \(2019\)](#). This strong correlation can be understood by looking at the following relation for an isothermal modified blackbody (see [Hayward et al., 2011](#)),

$$T_{\text{dust}} \propto \left(\frac{L_{\text{IR}}}{M_{\text{dust}}} \right)^{1/6} \quad (\text{for } \beta = 2). \quad (5.4)$$

SEDs can be qualitatively described by such a form. For main sequence galaxies, $L_{\text{IR}}/M_{\text{dust}}$ has

been found to be proportional to the sSFR (e. g. [Magdis et al., 2012](#); [Magnelli et al., 2014](#); [Ma et al., 2019](#)), implying an inverse correlation with λ_{peak} . We show on the figure that this matches our results well at high sSFR (the $T_{\text{peak}} \propto \text{sSFR}^{1/6}$ dashed line). This relation can also be used to understand the increasing dust temperatures with redshift (explicit redshift evolution is shown in Figure 5.10). [Lovell et al. \(2021a\)](#) has already shown that there is a systematic increase in the normalisation of the sSFR of the FLARES galaxies at constant stellar mass. This would imply that the redshift dependence of the dust temperature can be attributed to the increasing sSFR. We explore the evolution of λ_{peak} with different stages of galaxy star-formation activity in Appendix 5.D.

5.3.3.2 T_{SED}

To obtain the SED dust temperature we follow [Casey \(2012\)](#) by parameterising our galaxy SEDs using the sum of a single modified-blackbody and a mid-infrared powerlaw. The addition of the powerlaw to the functional form provides a better fit to the mid-infrared which is dominated by warm dust. Using this prescription, the luminosity at a rest-frame frequency ν can be written as

$$L_{\nu}(\nu) = N_{\text{bb}}f(\nu, \beta, T_{\text{SED}}) + N_{\text{pl}}(\nu_c/\nu)^{-\alpha} e^{-(\nu_c/\nu)^2}. \quad (5.5)$$

Here,

$$f(\nu, \beta, T) = \frac{(1 - \exp(-(\nu/\nu_1)^{\beta}))}{\exp(h\nu/(k_b T)) - 1} \nu^3, \quad (5.6)$$

$$N_{\text{pl}} = N_{\text{bb}}f(\nu_c, \beta, T_{\text{SED}}), \quad (5.7)$$

where h and k_b are the Planck's constant and the Boltzmann constant, respectively. The free parameters in the form are N_{pl} (normalisation factor), β (emissivity index), T_{SED} (SED dust temperature), ν_1 (frequency where optical depth is unity, usually taken as $\sim 100\mu\text{m}$ or 3THz in high-redshift studies) and α (mid-IR power-law slope). We adopt the same parameterisation of ν_c as given in [Casey \(2012\)](#) (or λ_c there). We use LMFIT ([Newville et al., 2014](#)), a non-Linear least-squares minimization and curve-fitting package in PYTHON, to fit this parametric form to the SKIRT SEDs and obtain T_{SED} . In the fitting, we impose the criteria that the dust temperature is higher than the CMB temperature at that redshift. It should also be noted that there is degeneracy between the dust temperature and the emissivity; a higher temperature can compensate for a lower value of the emissivity, and vice-versa. Thus when comparing to observations there is considerable maneuverability when choosing the values of T_{SED} and β , and thus deviations or agreement with the values can also be achieved based on the ranges being probed. In our case we

see that in general the SED temperature increases while the emissivity decreases with increasing redshift when they are both kept as free parameters.

The Rayleigh-Jeans (RJ) part of the SED can also be approximated by a generalised modified-blackbody function in the optically thin case (equation 2 in Casey, 2012) by

$$L_\nu(\nu) = A \frac{\nu^{3+\beta}}{\exp(h\nu/(k_b T_{\text{SED,RJ}})) - 1}, \quad (5.8)$$

where A is a normalisation constant and the rest of the terms are defined as before. In this case the free parameters are A , β (parameter search is confined to the range 1.5 – 2.5 to be consistent with high-redshift observational works referenced here) and $T_{\text{SED,RJ}}$. We refer to the SED dust temperature obtained from this functional form as $T_{\text{SED,RJ}}$. Similar to the fitting function in equation 5.5, a similar degeneracy exists here as well. This form is usually used by observational studies to derive the total infrared luminosity of galaxies (e.g. Knudsen et al., 2017; Hashimoto et al., 2019; Bouwens et al., 2020). In many cases where there is only a single detection in the dust-continuum, β and $T_{\text{SED,RJ}}$ are kept constant and a fit for the normalisation is obtained. It should also be noted that the SED dust temperature obtained from this form closely matches with the galaxy peak dust temperatures, while T_{SED} is typically higher than the peak dust temperature (see Figure 2 in Casey, 2012). From our analysis we also see that this form can lead to an overprediction of the obtained total infrared luminosity (median deviation of $\sim +17\%$ at $z = 5$ and lowering to $\lesssim 1\%$ by $z = 10$) compared to results using equation 5.5 or the true SED. This is seen when we use the full range of the dust SED. We have also tried to constrain our fits by only using wavelength ranges in the RJ tail. In this case, most fits underpredicted the total IR luminosity. Thus it is important to have some constraints at wavelengths short of the RJ tail to produce reliable SED temperature estimates that can retrieve the total IR luminosity.

5.3.3.3 Redshift evolution of dust temperatures

In Figure 5.10 we show the weighted median evolution of the different dust temperatures for the FLARES galaxies. We also show the 16-84th percentile spread of the values as well as the error on the median. Overplotted are several observational dust temperature values from high-redshift galaxies.

The figure clearly indicates that the median values of the dust temperatures consistently increase towards higher redshift, as expected since higher redshift galaxies are intensely star-forming (higher sSFR), which leads to higher UV emission resulting in warmer dust. There is also a spread of $\sim 10\text{K}$ in all the temperature values across the redshift range. T_{peak} increases from a median value of $\sim 40\text{K}$ at $z = 5$ to $\sim 70\text{K}$ at $z = 10$, which is higher than the increase in

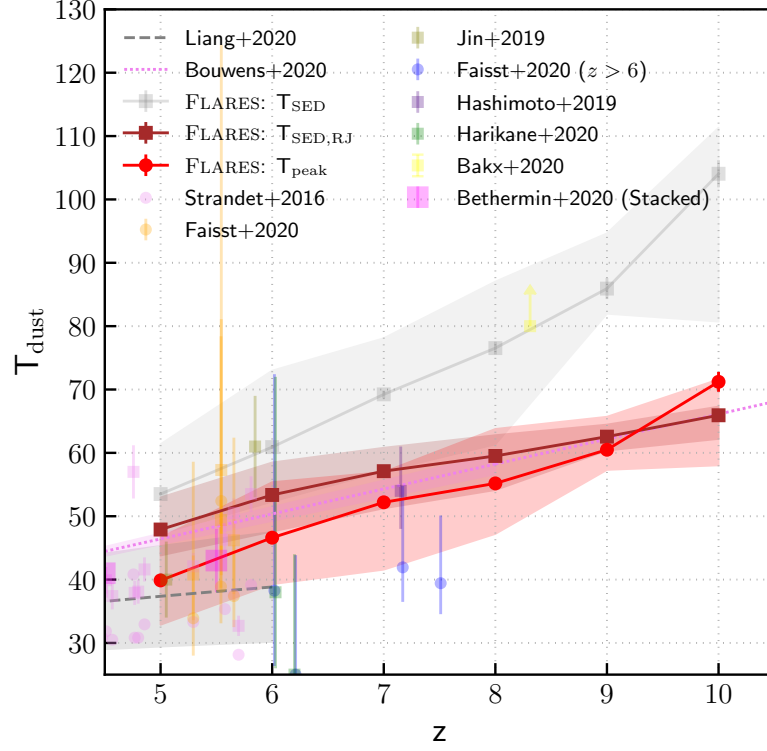


Figure 5.10: We show the evolution of the peak dust temperature (T_{peak} , red circles) and the SED dust temperature (brown squares for $T_{\text{SED,RJ}}$, grey squares for T_{SED}) from the FLARE simulation. The markers indicate the weighted median, the 16–84 percentile spread and the error on the median (the errorbars, negligible due to the high number counts) at $z \in [5, 10]$. Observational data from studies at high redshift (circle for T_{peak} and square for T_{SED} or $T_{\text{SED,RJ}}$ values); included are data from Strandet et al. (2016, both T_{peak} and $T_{\text{SED,RJ}}$), Jin et al. (2019, T_{SED} calculated from optically thick MBB), Faisst et al. (2020b, both T_{peak} and T_{SED}), also included are the remeasured T_{peak} values from Knudsen et al. 2017; Hashimoto et al. 2019), Hashimoto et al. (2019, $T_{\text{SED,RJ}}$), Harikane et al. (2020, $T_{\text{SED,RJ}}$) Béthermin et al. (2020, $T_{\text{SED,RJ}}$ obtained from stacked galaxies with $\text{SFR} \geq 10 M_{\odot}/\text{yr}$) and Bakx et al. (2020, $T_{\text{SED,RJ}}$, lower limit) are plotted. We also show the fit functions for T_{peak} from Liang et al. (2019, dashed line, with the shaded region showing the spread) and for $T_{\text{SED,RJ}}$ from Bouwens et al. (2020, dotted line) (see text for more details).

the CMB temperature across this redshift. Using our sample of massive galaxies, we fit a linear relation to the median redshift evolution of T_{peak} and obtain

$$T_{\text{peak}}/\text{K} = (40.08 \pm 0.15) + (5.75 \pm 0.13)(z - 5). \quad (5.9)$$

The relation has a higher slope compared to the one in Schreiber et al. (2018) (slope of 4.60 ± 0.35) obtained for an observational sample using stacked SEDs of main-sequence galaxies at $z \leq 4$. This indicates that towards higher redshift in the EoR, the evolution of T_{peak} is stronger for the massive galaxies in FLAREs than their low redshift relation.

In Figure 5.10, we also compare our values to other theoretical and observational results at similar redshifts. We compare to observational values from Strandet et al. (2016); Faisst et al. (2020b), with most values from FLAREs in good agreement or otherwise within the constraints.

There are a few galaxies in [Strandet et al. \(2016\)](#) which show slightly colder T_{peak} values, in tension with our predictions. FLARES fails in this case to be fairly representative of such cold dusty galaxies. It can be seen that our T_{peak} values are offset from the relation obtained from the MASSIVEFIRE simulations in [Liang et al. \(2019\)](#) for $2 \leq z \leq 6$. The results are in agreement in the region for which the fit was obtained, but would be an overprediction on extrapolation. This difference could be due to the smaller sample size (29 massive galaxies in total) as well as the use of a higher dust-to-metal (DTM=0.4, increasing the optical depth) ratio in that study. Our result is also very similar to the recent values from the ILLUSTRIS-TNG ([Shen et al., 2021](#)) suite of simulations using SKIRT, with their reported median T_{peak} being slightly higher at $z = 4, 6, 8$.

We can compare our T_{SED} values to the ones in [Faisst et al. \(2020b\)](#), since they use the same fitting function as in Equation 5.5 (in their work α is fixed at 2.0, since the SED is not constrained blueward of rest-frame $\sim 110\mu\text{m}$). The values obtained in that study provide a very reasonable match to our constraints within the median spread, while the other observational results are lower by $\gtrsim 10\text{K}$. The measurement from [Jin et al. \(2019\)](#) uses an optically thick MBB that gives dust SED temperatures that are very similar to the ones obtained from equation 5.5 (see Figure 2 in [Casey, 2012](#)). One of the values is in very good agreement with our measures for T_{SED} , while the other galaxy at $z \sim 5$ in the work has a very cold dust SED temperature.

For all the other measurements it would be fairer to compare the values to $T_{\text{SED,RJ}}$, as equation 5.8 is used in many of the high-redshift studies to fit for the observed fluxes/luminosities. As such our predicted values provide a reasonable match to the observational data from [Strandet et al. \(2016\)](#), [Hashimoto et al. \(2019\)](#) and [B  thermin et al. \(2020\)](#), measures the temperature using stacked galaxies with $\text{SFR} \geq 10 M_{\odot}/\text{yr}$. There are a few exceptions in the observational data that deviate strongly from our predictions. For example, [Harikane et al. \(2020\)](#) fit an optically thin MBB to their galaxies at $z \sim 5$, and find very cold dust SED temperatures. However, in the case of [Bakx et al. \(2020\)](#) the lower limit they provide is very high compared to our predictions, which could be reconciled if using a much high emissivity index. Some of the large dispersions seen in the dust temperature measures in observations point towards either a wide range of values existing in the diverse populations of galaxies in the early Universe, or an indication of more varied dust grain properties such as their size, shape, and composition that is not captured in the models being employed to study them. The [Bouwens et al. \(2020\)](#) fit to the dust SED temperatures (with some of the [Faisst et al. 2020b](#) peak dust temperature values also being used), which were used in their modified blackbody fitting, does have a few values at $z \geq 5$, however the majority of the constraints are from lower redshift (see their Figure 1). The FLARES dust SED temperatures (referring to $T_{\text{SED,RJ}}$) have slightly higher normalisation compared to their fit, but are in reas-

onable agreement within the 16-84th percentile spread. Similar to our results, the T_{SED} values in Shen et al. (2021) are also consistently higher than the observationally quoted SED temperatures, with their median being slightly higher than our values.

It should also be noted that the plotted values and the spread are weighted based on which overdensity region they are from, and thus contribution from the extreme overdensities will be down-weighted. Even though they are rare, all the observational values outside the $1-\sigma$ scatter are well within the maximum and minimum values of the seen dust temperatures across redshifts in FLARES, except for the lower limit measurement from Bakx et al. (2020).

5.4 Conclusions

We have presented the dust SED properties of galaxies in FLARES, a suite of zoom simulations that uses the EAGLE (Schaye et al., 2015; Crain et al., 2015) physics to probe a range of overdensities in the EoR. We select massive galaxies ($\gtrsim 10^9 M_\odot$) in the simulation to make a comprehensive statistical study of galaxies that are accessible to current telescopes. These galaxies are post-processed with the radiative transfer code SKIRT (Camps & Baes, 2015, 2020) to generate their full SEDs. The dust-to-metal ratios were derived from the fitting function from the dust model implemented in the L-Galaxies SAM (Vijayan et al., 2019). We do not calibrate any of the parameters in SKIRT to produce the SEDs.

Our main findings are as follows:

1. The predicted UV LF is in agreement with available observational data. We also compare the IR LF to the observations, finding good agreement at $z = 5, 6$ for luminosities $< 10^{12} L_\odot$. We underestimate the number densities of the most IR luminous galaxies at $z = 5$. We attribute this mainly to the lack of high star formation rates in the massive galaxies in the simulation, similar to what previous FLARES or EAGLE studies have shown (e.g. Katsianis et al., 2017; Baes et al., 2020; Lovell et al., 2021a). We also underpredict the number of luminous rest-frame 250 μm galaxies. However, the observations Koprowski et al. (2017); Gruppioni et al. (2020) show discrepancies among each other. The extreme IR objects are biased towards the highest matter overdensities.
2. The FLARES IRX- β relation for $5 \leq z \leq 8$ is consistent with the starburst relation (e.g. Meurer et al., 1999; Reddy et al., 2015) from local redshifts. We see a shift towards the SMC relation (Pettini et al., 1998) for $z > 8$. We see that it is predominantly the lower-mass ($\lesssim 10^9 M_\odot$) galaxies that deviate towards this relation. Also galaxies with low-sSFR lie

further away from these empirical relations. We see a good match with current available observations, missing a few low-IRX high- β galaxies.

3. The IRX shows a gradual increase with stellar mass, showing a flattening at high-stellar masses ($\sim 10^{10} M_{\odot}$). We do not see any evolution in the normalisation of the median relation with redshift.
4. We look at the evolution of the peak of the IR emission (λ_{peak}) with redshift on properties like the galaxy stellar mass, total IR luminosity and the sSFR. We see flattening of λ_{peak} at high IR luminosity ($\sim 10^{11} L_{\odot}$). The λ_{peak} (T_{peak}) - L_{IR} relation is offset from the observed local relation (Casey et al., 2018) to lower (higher) values. λ_{peak} strongly correlates with the galaxy sSFR.
5. Luminosity-weighted dust temperatures (peak dust-temperature: T_{peak} , SED temperature fit from mid-IR powerlaw+MBB: T_{SED} and SED temperature fit from optically-thin MBB: $T_{\text{SED,RJ}}$) increase with increasing redshift. We find that, for the massive galaxies in FLARES, the evolution of T_{peak} with redshift is stronger than the low-redshift relation obtained from observational (Schreiber et al., 2018) and other theoretical (Liang et al., 2019) studies.
6. The SED temperatures (T_{SED} and $T_{\text{SED,RJ}}$) are mostly in agreement with the observational values. However we find a lack of extremely cold temperatures seen in some observations (Strandet et al., 2016; Jin et al., 2019).

Future observations from many of the planned surveys and observations on *ALMA*, *JWST*, *Roman Space Telescope*, *Euclid*, as well as future IR missions sampling more of the SED will be able to put better constraints on these dust driven properties. In a future work we will explore the dust-continuum sizes of these galaxies.

Through this study as well as previous other works referenced here, there is some evidence in favour of the EAGLE physics model requiring higher star-formation rates to match some of the observations at high-redshift like the UV LF, IR LF or the sub-mm number counts. However, reconciliation of such limitations must be achieved without losing some of the remarkable successes of the model across the low-redshift Universe. The strive to succeed in this extremely non-trivial challenge has been the goal of all theoretical studies of galaxy formation and evolution. The high-redshift Universe is a regime where EAGLE as well other periodic boxes have not been well studied due to its lack of massive galaxies. Studies with FLARES allows for a statistical exploration of this regime due its novel re-simulation strategy targeting massive overdensities. This will inevitably help to improve theoretical models of galaxy formation and evolution in

terms of providing insights into the different feedback mechanisms as well as star formation recipes implemented.

5.A Convergence Tests

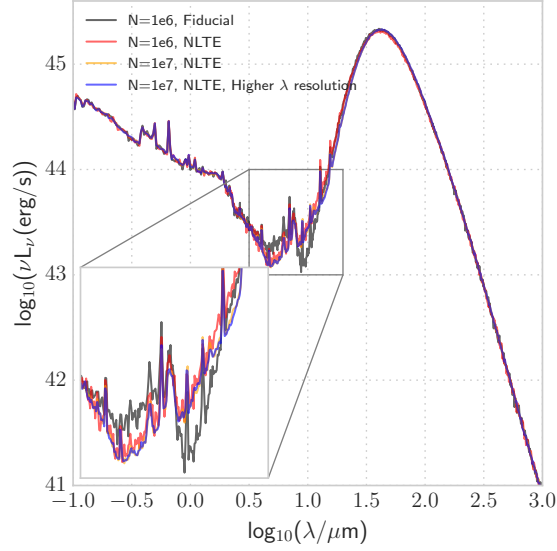


Figure 5.11: Shows the SED of a galaxy at $z = 8$ for a few different configurations of the SKIRT code. Shown in inset, the mid-IR region. The galaxy has a total IR luminosity of $\sim 10^{11.8} L_{\odot}$ in all the plotted configurations.

In this section we will test the convergence of our SEDs in relation to some of our SKIRT parameter choices. We will first explore how the choice of an increase in the photon number, increasing resolution of our radiation and dust wavelength grid, and adding Stochastic (Non Local

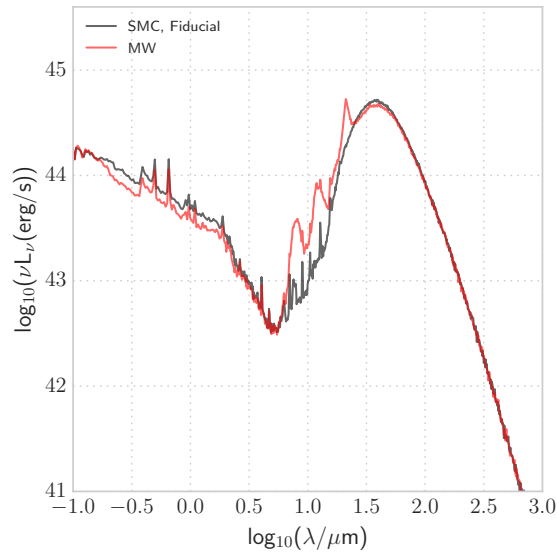


Figure 5.12: Shows the SED of a galaxy at $z = 8$ for the SMC and Milky Way grain distribution.

Thermal Equilibrium, NLTE) heating to the set-up can change our results. In Figure 5.11, we plot the SED of a galaxy at $z = 8$ for these different configurations. Our higher photon number test runs have 10^7 (10 times higher compared to the default run) photons per radiation field wavelength grid, while for the increased wavelength grid set-up we have doubled the number of bins from our default configuration. The total IR luminosity only changes by ~ 0.01 dex for the galaxy between these choices, while there is even smaller effect in the UV and optical part of the SED. We have checked this for a few other galaxies and find that the changes are similar. There are no noticeable dramatic changes.

We also change our dust grain distribution choice from SMC like to Milky Way (MW) like in Figure 5.12. The MW configuration has Poly-Aromatic Hydrocarbons (PAHs) included in the dust grain distribution, which will have an effect on the mid-IR range of the SED. There is a clear indication of an absorption feature near 2175\AA , due to the bump in the extinction curve for the MW dust distribution. This ultimately leads to β values that are always negative. It can be seen from the figure that there is also stronger extinction at wavelengths short of the mid-IR for the MW type, with higher emission in the mid-IR compared to SMC type. Next generation instruments that can scan the mid-IR SED at high-redshift are needed to put constraints in this regime, and to aid our understanding of emission from PAHs. The change to MW type grain has negligible effect on where the peak of the IR emission is. The total IR luminosity also sees negligible change towards higher values. However, the value of β is affected and thus can drive changes in the IRX- β plane by making the β values more negative.

5.B Effect of AGN on the dust SED

In order to understand the effect AGN have on the observed galaxy SEDs, we perform a simple analysis. For this purpose, we obtain the intrinsic bolometric luminosity of the SMBH in our galaxy sample, and add it to the total IR luminosity of our galaxies. This would represent an upper limit on the total IR luminosity that the galaxy can have from AGN contribution. The SMBH bolometric luminosity is calculated using

$$L_{\text{BH,bol}} = \eta \frac{dM_{\bullet}}{dt} c^2, \quad (5.10)$$

where dM_{\bullet}/dt is the accretion rate and η is the efficiency, assumed to be 0.1.

In Figure 5.13, we plot the IR LF with the AGN bolometric luminosity added to the FLARES galaxies for $z \in [5, 10]$. For comparison we also show the IR LF (which only includes stellar reprocessed dust emission) plotted in Figure 5.4. We do not plot any of the observations that

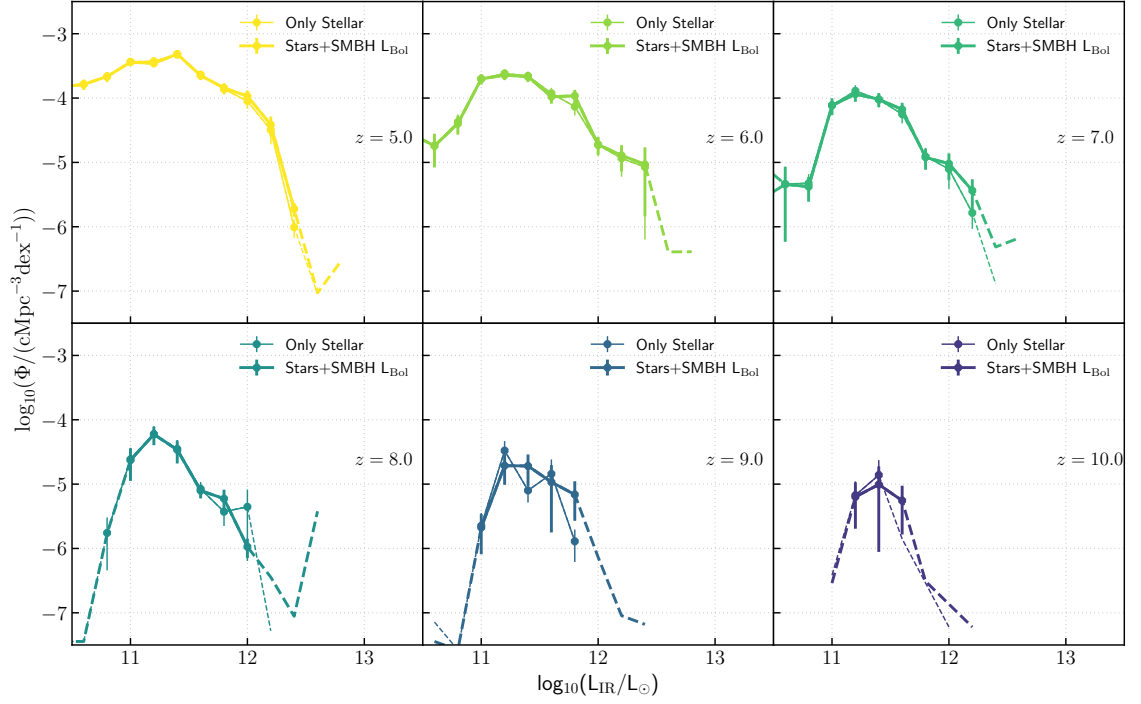


Figure 5.13: Same as Figure 5.4, now showing how the total infrared luminosity function changes, if all the energy from the SMBH accretion, as described in Equation 5.10 went into the infrared. This is shown with thicker lines, while the fiducial IR LF is shown by lighter lines.

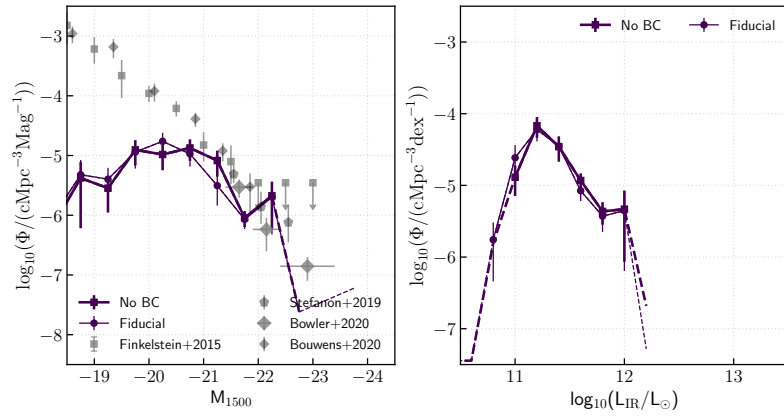


Figure 5.14: Plot compares the UV (left panel) and IR (right panel) LF with (thin line) and without (thick line) the birth cloud attenuation for $z = 8$. Observational data, same as in Figure 5.3, is also plotted in the left panel.

were shown in Figure 5.4. We can see that there is a small change at the very bright end of the function. However this increase is not enough to reconcile the relation with the observational data. Thus, as explained in §5.3.1, the main driver of the difference can be attributed to the lack of more intense star formation activity in massive/bright galaxies in the model.

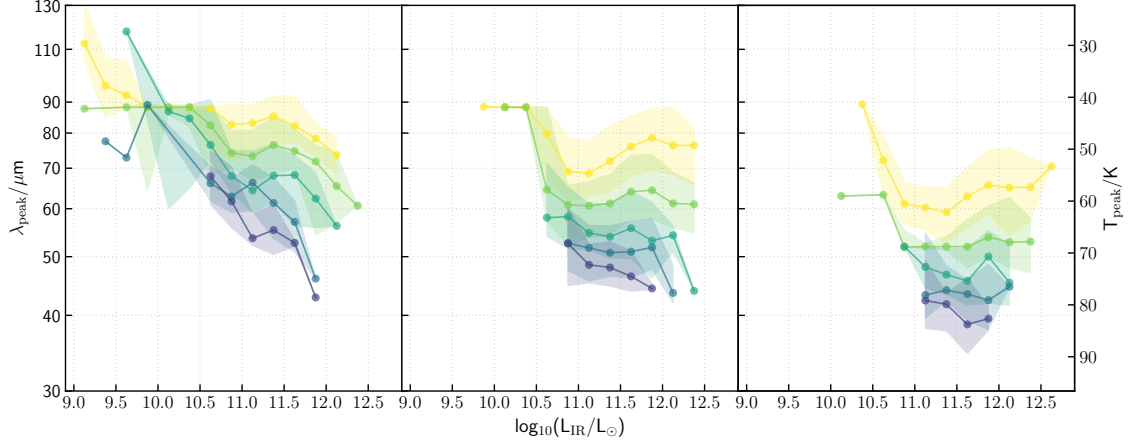


Figure 5.15: We show the variation of λ_{peak} (corresponding T_{peak} values is shown on the right y-axis) with the IR luminosity for $z \in [5, 9]$. The panels represent galaxies with suppressed star-formation (left), galaxies on the main-sequence (middle) and starburst galaxies (right).

5.C Excluding birth cloud emission

To see the effect of birth cloud emission in the model, we ran SKIRT for galaxies only at $z = 8$ (in order to drastically reduce the computational costs, since there are only fewer galaxies at $z = 8$ compared to lower redshifts), by treating the young star-forming particles as regular star particles, without the added dust extinction implemented in the MAPPINGS III template due to the birth cloud. We now model both these radiation sources using BPASS, ignoring this extra extinction. We plot the UV and IR LF of the results in Figure 5.14, as well as compare to our fiducial set-up. It can be seen that the UV LF is within the scatter at each bin. Similarly, in case of the IR LF at $z = 8$, there is only very negligible change. The small impact on both these functions is due to the low metallicity of these systems, owing to them being at extremely high-redshift. It is expected that with the increase in metallicity of systems at lower redshifts, birth cloud attenuation will have more of an influence.

5.D T_{peak} and galaxy main-sequence

To better understand the relation between λ_{peak} and the sSFR, we separate the FLARES galaxies into 3 groups, based on an evolving piecewise fit to the stellar mass-SFR relation presented in Lovell et al. (2021a, see §3.4 in the work, equations 11 and 12). The groups have been classified based on their deviation from the piecewise fit. The 3 groups have been plotted in Figure 5.15, with the median λ_{peak} as a function of the IR luminosity. These groups are galaxies

- below 1σ (left panel in Figure 5.15) from the fit, which can be classified to include the green

valley and the passive galaxies,

- within 1σ (middle panel in Figure 5.15) of the fit relation on either sides, termed the main-sequence, and
- above 1σ (right panel in Figure 5.15) of the fit, which can be termed as starbursts.

We only show in Figure 5.15 the relation for $z \in [5, 9]$, since the star-forming sequence fit could not be constrained properly at $z = 10$ (see Lovell et al., 2021a).

It can be seen from Figure 5.15 that the shape of the relation between λ_{peak} and L_{IR} is different for the 3 groups. Galaxies in the green valley/passive regime (left panel) show a consistent decrease (increase) in λ_{peak} (T_{peak}) with L_{IR} . This is mainly due to the smaller dust content within these galaxies and thus there is a direct correlation between the increase in dust-temperature and L_{IR} . The other two groups exhibit a flat relationship with L_{IR} , due to their high dust content and thus the hot dust being optically thick, similar to that seen in some observations at high-redshift (e.g. Cortzen et al., 2020). The starburst galaxies have a lower median λ_{peak} due to their higher sSFRs. At $z = 5$, towards high L_{IR} there is also a hint of increasing λ_{peak} values, indicating the rapid build up of dust in these extreme objects.

5.E Comparison with line-of-sight model

In this section we will compare the UV luminosity obtained from our SKIRT modelling here to the line-of-sight (LOS) dust model we implemented on the same galaxies in Chapter 4. In that work we assumed a dust attenuation curve and modelled the dust attenuation parameters for the old stars and the young stars to find a good match to the $z = 5$ UV LF and the UV- β relation, as well as observations of the [OIII]+H β EW relations at $z = 8$. This method is computationally much faster, and allows for more flexibility in the modelling, allowing you to explore changes and their effects more easily. However, it can not treat certain phenomena, such as the scattering of light away from the LOS or dust self-absorption, as these processes are dependent on the chosen extinction curve.

We compare the UV luminosity of the galaxies selected in this work, using the two dust models in Figure 5.16, for $z \in [5, 10]$. We can see that the values obtained in this work are systematically lower than the ones from the LOS model by ~ 0.4 dex. This is mainly due to the lower dust optical depth (parameterised by the κ parameters) along the LOS adopted in that study compared to this work. This by construction matches the observations that the model was

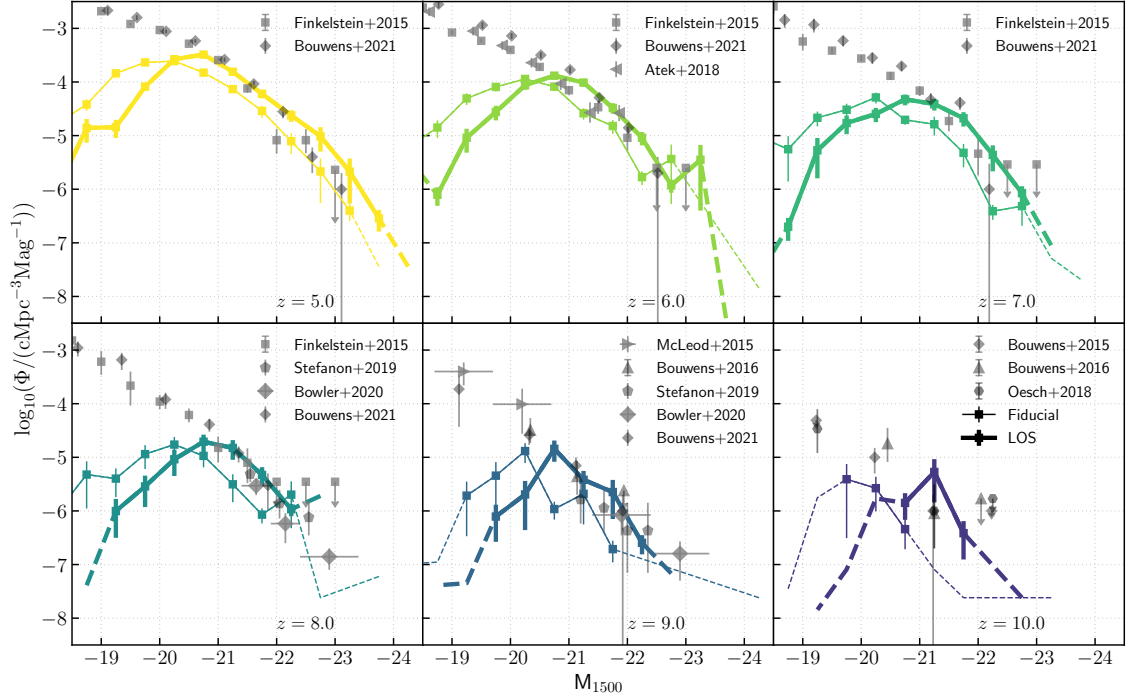


Figure 5.16: Same as Figure 5.3, but here we also include the UV LF obtained from the LOS dust extinction model (thicker lines) implemented in Chapter 4 on the FLARES galaxies.

calibrated for. This implies that to match the SKIRT results the dust-optical depths should be higher, i. e. the value of κ_{ISM} (keeping κ_{BC} constant) parameter should be higher to reproduce the SKIRT result, and still broadly match the UV continuum as well as some of the line luminosity relations.

We have already explained one of the reasons for the slightly lower number densities in §5.3.1 for the non-calibrated model presented here.

CONCLUSIONS

6

In this thesis, I have used simulations of galaxy formation and evolution, both Semi-Analytical Models (SAMs) and hydrodynamical simulations to study the evolution of dust in galaxies and its effect in observing the early Universe. As seen through the chapters in this thesis as well as the wealth of observational data available, dust plays a major role, even at these high redshifts ($z \geq 5$) to influence the observational interpretations. In this thesis I have tried to provide simple models as well as complicated radiative transfer approach to incorporate the effect of dust. Below, I briefly summarise the main conclusions derived from the works.

In Chapter 2, we build a simple model for dust production and destruction into the L-GALAXIES SAM version of [Henriques et al. \(2015\)](#). The model includes prescriptions for dust production from AGB stars, supernovae ejecta and grain growth as well as destruction due to supernovae explosions and incorporation into the hot halo. The grain growth mechanism is similar to the one used in [Popping et al. \(2017a\)](#), but we follow the dust content separately in molecular clouds and the diffuse medium. We see an evolution in the dust-to-metal (DTM) ratio of galaxies through redshift, which can be parameterised using the gas-phase metallicity and age of the stellar population. We find that grain growth is the dominant dust production mechanism at $z < 8$, with a significant number of galaxies at $z = 6$ showing a transition from dust injection through type II supernovae to production via grain growth.

In Chapter 3, we introduce our suite of simulations termed First Light And Reionisation Epoch Simulations (FLARES) to study the Epoch of Reionisation (EoR). FLARES follows the evolution a wide range of overdensities selected from a $(3.2 \text{ Gpc})^3$ dark matter only volume, resimulated using the well studied EAGLE model ([Schaye et al., 2015](#); [Crain et al., 2015](#)). The access to some of the largest overdensities in the early Universe enables us to explore the properties of galaxies that will be studied using the next generation of surveys and telescopes. Using a

novel weighting scheme we can combine the regions to produce composite distribution functions. We find that the galaxy stellar mass function and the star formation rate function of the FLARES galaxies follow a double Schechter shape, with the most massive and star-forming galaxies biased to the most overdense regions. We find excellent matches to the stellar mass and SFR function against observations, which gave us confidence in using the simulations to derive the photometric properties, the main content of the next chapter.

In Chapter 4, we apply a simple line-of-sight dust extinction model to the galaxies in FLARES, similar to the work of Wilkins et al. (2016c), to produce the UV to near-IR SEDs. We improve upon that model by splitting the extinction into contributions from the birth cloud of young stars and the ISM as well as include a model of nebular emission from Wilkins et al. (2020). The birth cloud extinction is proportional to the metallicity of the star particle, while the extinction due to ISM is proportional to the metal density along the line of sight multiplied by the DTM ratio calculated from the DTM fitting equation from the L-GALAXIES model. A novel feature about this work is the use of a varying DTM ratio for different galaxies. We see excellent agreement of the derived photometric properties at different redshifts with observations. We predict the amount of obscured and unobscured star formation rate density to have comparable values by $z \sim 6$, with obscured star formation dominating below this redshift. We also do not see any significant variation in the shape of the luminosity function with different environments with the brightest galaxies biased towards the most overdense regions. Our datasets are publicly available and can be used to make predictions for the next generation surveys.

In Chapter 5, we post-process the most massive galaxies ($\gtrsim 10^9 M_\odot$) in FLARES with the radiative transfer code SKIRT (Camps & Baes, 2015) to produce the full SEDs. We assign the galaxy DTM ratio from the L-GALAXIES fit function. We compare the UV, IR and $250\mu\text{m}$ LF against observations, finding reasonable agreement. We underpredict the number densities of bright IR galaxies at $z = 5$, attributing it to the lower SFR seen in the EAGLE model or the need of top-heavy IMFs in the most star-forming galaxies. We see that the FLARES $\text{IRX-}\beta$ relation is consistent with the local starburst relation. We also see an evolution of the peak dust temperature (T_{peak}) with the IR luminosity. T_{peak} strongly correlates with the galaxy sSFR. All the luminosity-weighted temperatures considered in the study increases with increasing redshift. We find a stronger evolution of T_{peak} with redshift compared to other low-redshift studies.

The body of work presented above lies in a regime whose predictions are testable with the current and upcoming facilities such as ALMA, AtLAST, Euclid, JWST, Roman, etc. We have briefly touched upon these in the different chapters such as the dust masses and dust temperature at high redshift, which many ALMA current programmes are exploring, enabling us to under-

stand the characteristics of dust at these redshifts. *JWST* will complement these observations in the UV and optical, providing data on the star formation rate, metallicity, dust attenuation and ionisation in the most massive galaxies, the characteristics of which can be understood in terms of our modelling framework. *Euclid* and *Roman* will be able to add better constraints on the bright-end of the UV luminosity function with its deep and wide survey capabilities. Many key predictions from our modelling such as the redshift evolution of the DTM ratio or the existence of extreme UV-bright galaxies with lower attenuation or the amount of obscured SFR at high-redshift or the IR optical depth of massive/IR-bright galaxies or the dust temperature evolution will be probed with more statistical power in the coming years. The number density of passive galaxies is another realm to test the predictions. These results will help in the advancement of theoretical simulations at high-redshift providing us insight into the implemented feedback and star formation recipes.

6.1 Future Works

One of the future plans we have is to import the dust production/destruction model into the latest version of L-GALAXIES (Henriques et al., 2020), where the galaxy components have been split into 12 concentric rings. It would be a great avenue to explore the dust radial profiles of galaxies. We also plan to improve the framework by including a model to incorporate dust in the hot halo as well.

A project that I am currently working on is to look at the radial and orientation variation in the UV attenuation. The former would allow us to see how much the assumption of a single or dual (like the Charlot & Fall 2000 implementation) attenuation across stars impacts different values derived from SED modelling, while the latter would allow us to understand the scatter in the attenuation from galaxy orientation. We see huge scatter in the attenuation for the different star particles within a galaxy, meaning global attenuation values are misrepresentative. We also see that there is a scatter in the galaxy UV attenuation due to line-of-sight effects, which can be quantified using the morphology of a galaxy, primarily of how disk-like it is and the dust surface density, which is shown in Figure 6.1. We parameterise the diskiness of a galaxy by the use of $\kappa_{\text{corotation}}$ (κ_{co} , with values between 0-1, with greater values indicating higher ordered rotation or being more disk-like). This value in a nutshell quantifies the energy invested in ordered corotation (see Correa et al. 2017 for more details). This is calculated using only star particles that follow the direction of rotation of the galaxy. Correa et al. (2017) referred to galaxies with $\kappa_{\text{co}} < 0.4$ as ellipticals and galaxies with $\kappa_{\text{co}} > 0.4$ as discs. Here, we do not need to make any explicit

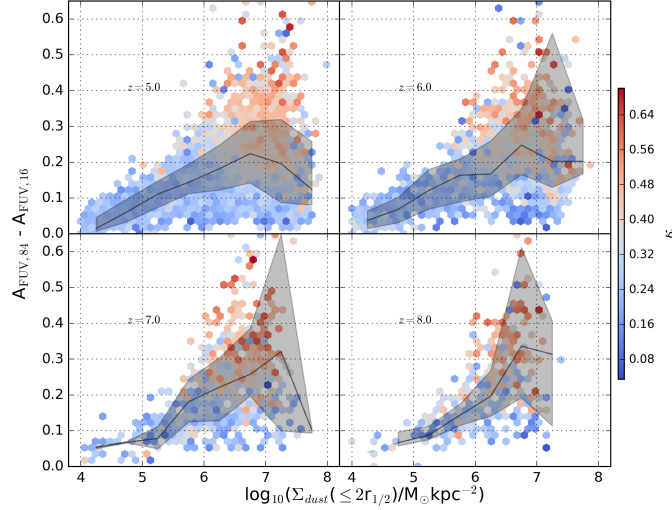


Figure 6.1: Shows the spread in the UV attenuation for different lines-of-sight as a function of the dust surface density (Σ) inside twice the stellar half-mass radius for galaxies with more than 1000 star particles for $z \in [5, 8]$. The hexbins are coloured by κ_{CO} , a parameter which quantifies how disk-like a galaxy is.

distinction between galaxies to be ellipticals or discs, and can use this value to quantify how disk-like a galaxy is. We plan to use machine learning techniques to quantify feature importance as well.

While generating the full SEDs using SKIRT, we also generated data cubes containing pixel information of the galaxies in the UV and IR. We plan to use this to quantify the dust-continuum sizes of the galaxies at high-redshift. Some recent studies see offsets on the scale $> 2 - 3$ kpc (e.g. [Bowler et al., 2018](#)); we can use FLARES to see if simulations can predict these objects as well as quantify if they are from single objects or adjacent substructure. The results from these would have implications on star formation rates in the early Universe and help in our understanding of various SFR calibration relations, such as the IRX- β relation (which in the local Universe has been verified to vary spatially in resolved studies).

There are also plans for the next phase of the FLARE simulations termed FLARES-II, where we run new set of regions (10 times more) selected from a much larger volume (current plan is a 6.4 Gpc box) using the latest physics from the EAGLE-SWIFT simulations. This is at the calibration stage and will start its science runs in the near future.

FURTHER ACKNOWLEDGMENTS



For Chapter 2, the authors would like to thank the anonymous referee for a constructive and detailed report that has significantly enhanced the strength of the work. Thanks also to Gergö Popping and Phil Wiseman for useful discussions during the undertaking of this work and Pierre Ocvirk for his helpful correspondence while using our fitting function. Much of the data analysis was undertaken on the APOLLO cluster at the University of Sussex.

For Chapter 3, the authors wish to thank the anonymous referee for detailed comments and suggestions that improved this work. We also thank Scott Kay and Adrian Jenkins in particular for their invaluable help getting up and running with the EAGLE resimulation code. Thanks also to Rob Crain, Rachana Bhatawdekar, Daniel Ceverino and Kristian Finlator for helpful suggestions and discussions. Finally, we thank the EAGLE team for their efforts in developing the EAGLE simulation code.

For Chapter 4, we wish to thank the anonymous referee for detailed comments and suggestions that improved this paper. The authors also thank Desika Narayanan for providing the extinction curve used in Narayanan et al. (2018), and Rebecca Bowler for providing the DPL fit parameters.

For Chapter 5, the authors thank Mark Sargent and Rebecca Bowler for helpful discussions. We also thank the ALPINE-ALMA collaboration for making their data products public. We also thank Jorge Zavala and Yueying Ni for helpful correspondences.

Chapter 3, 4 and 5 used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, ST/R002371/1 and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1.

DiRAC is part of the National e-Infrastructure.

I wish to acknowledge the following open source software packages used in the analysis: NUMPY ([Harris et al., 2020](#)), SCIPY ([Virtanen et al., 2020](#)), ASTROPY ([Astropy Collaboration et al., 2013, 2018](#)), MATPLOTLIB ([Hunter, 2007](#)) and WebPlotDigitizer ([Rohatgi, 2020](#)). I acknowledge the support of my PhD studentship from UK STFC DISCnet.

A significant part of the work presented in this thesis was done during the global Covid-19 pandemic. I would like to acknowledge the tireless efforts of the essential workers, who did not have the safety of working from their homes as well as our privileged positions that provided us the opportunity to pursue research in these tough times, while many members of our global society cannot afford to do that. I would also like to offer my deepest gratitude to everyone behind the development of various Covid-19 vaccines.

BIBLIOGRAPHY

- Anders P., Fritze-v. Alvensleben U., 2003, *A&A*, 401, 1063
- Angulo R. E., Hilbert S., 2015, *MNRAS*, 448, 364
- Angulo R. E., White S. D. M., 2010, *MNRAS*, 154, 143
- Aoyama S., Hou K.-C., Shimizu I., Hirashita H., Todoroki K., Choi J.-H., Nagamine K., 2017, *MNRAS*, 466, 105
- Arrigoni M., Trager S. C., Somerville R. S., Gibson B. K., 2010, *MNRAS*, 402, 173
- Ashby M. L. N., et al., 2013, *ApJ*, 769, 80
- Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, *ARA&A*, 47, 481
- Astropy Collaboration et al., 2013, *A&A*, 558, A33
- Astropy Collaboration et al., 2018, *AJ*, 156, 123
- Atek H., Richard J., Kneib J.-P., Schaerer D., 2018, *MNRAS*, 479, 5184
- Baes M., Camps P., 2015, *Astronomy and Computing*, 12, 33
- Baes M., et al., 2020, *MNRAS*, 494, 2912
- Bahé Y. M., et al., 2016, *MNRAS*, 456, 1115
- Bahé Y. M., et al., 2017, *MNRAS*, 470, 4186
- Bakx T. J. L. C., et al., 2020, *MNRAS*, 493, 4294
- Baldry I. K., Glazebrook K., Driver S. P., 2008, *MNRAS*, 388, 945

- Barisic I., et al., 2017, *ApJ*, 845, 41
- Barnes J., Hut P., 1986, *Nature*, 324, 446
- Barnes D. J., Kay S. T., Henson M. A., McCarthy I. G., Schaye J., Jenkins A., 2017a, *MNRAS*, 465, 213
- Barnes D. J., et al., 2017b, *MNRAS*, 471, 1088
- Barrow K. S. S., Wise J. H., Norman M. L., O'Shea B. W., Xu H., 2017, *MNRAS*, 469, 4863
- Baugh C. M., 2006, *Reports on Progress in Physics*, 69, 3101
- Beckwith S. V. W., et al., 2006, *AJ*, 132, 1729
- Begelman M. C., Volonteri M., Rees M. J., 2006, *MNRAS*, 370, 289
- Bekki K., 2013, *MNRAS*, 432, 2298
- Benitez-Llambay A., 2015, py-sphviewer: Py-SPHViewer v1.0.0, doi:10.5281/zenodo.21703, <http://dx.doi.org/10.5281/zenodo.21703>
- Benson A. J., 2010, *Phys. Rep.*, 495, 33
- Bernstein R. A., Freedman W. L., Madore B. F., 2002, *ApJ*, 571, 56
- B  thermin M., et al., 2020, *A&A*, 643, A2
- Bhatawdekar R., Conselice C. J., 2021, *ApJ*, 909, 144
- Bhatawdekar R., Conselice C. J., Margalef-Bentabol B., Duncan K., 2019, *MNRAS*, 486, 3805
- Blitz L., Rosolowsky E., 2006, *ApJ*, 650, 933
- Bonafede A., Dolag K., Stasyszyn F., Murante G., Borgani S., 2011, *MNRAS*, 418, 2234
- Bondi H., Hoyle F., 1944, *MNRAS*, 104, 273
- Boselli A., et al., 2010, *PASP*, 122, 261
- Bouwens R. J., Illingworth G. D., Blakeslee J. P., Franx M., 2006, *ApJ*, 653, 53
- Bouwens R. J., Illingworth G. D., Franx M., Ford H., 2008, *ApJ*, 686, 230
- Bouwens R. J., et al., 2012, *ApJ*, 754, 83
- Bouwens R. J., et al., 2014, *ApJ*, 793, 115

- Bouwens R. J., et al., 2015, *ApJ*, **803**, 34
- Bouwens R. J., et al., 2016, *ApJ*, **830**, 67
- Bouwens R. J., Oesch P. A., Illingworth G. D., Ellis R. S., Stefanon M., 2017, *ApJ*, **843**, 129
- Bouwens R. J., Stefanon M., Oesch P. A., Illingworth G. D., Nanayakkara T., Roberts-Borsani G., Labbé I., Smit R., 2019, *ApJ*, **880**, 25
- Bouwens R., et al., 2020, *ApJ*, **902**, 112
- Bouwens R. J., et al., 2021, *AJ*, **162**, 47
- Bower R. G., Schaye J., Frenk C. S., Theuns T., Schaller M., Crain R. A., McAlpine S., 2017, *MNRAS*, **465**, 32
- Bowler R. A. A., et al., 2014, *MNRAS*, **440**, 2810
- Bowler R. A. A., et al., 2015, *MNRAS*, **452**, 1817
- Bowler R. A. A., Dunlop J. S., McLure R. J., McLeod D. J., 2017, *MNRAS*, **466**, 3612
- Bowler R. A. A., Bourne N., Dunlop J. S., McLure R. J., McLeod D. J., 2018, *MNRAS*, **481**, 1631
- Bowler R. A. A., Jarvis M. J., Dunlop J. S., McLure R. J., McLeod D. J., Adams N. J., Milvang-Jensen B., McCracken H. J., 2020, *MNRAS*, **493**, 2059
- Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, *MNRAS*, **398**, 1150
- Bridge J. S., et al., 2019, *ApJ*, **882**, 42
- Bromm V., 2013, *Reports on Progress in Physics*, **76**, 112901
- Bromm V., Larson R. B., 2004, *ARA&A*, **42**, 79
- Bromm V., Loeb A., 2003, *ApJ*, **596**, 34
- Bruzual G., Charlot S., 2003, *MNRAS*, **344**, 1000
- Bunker A. J., Stanway E. R., Ellis R. S., McMahon R. G., 2004, *MNRAS*, **355**, 374
- Bunker A. J., et al., 2010, *MNRAS*, **409**, 855
- Burnham A. D., et al., 2021, *ApJ*, **910**, 89
- Byler N., Dalcanton J. J., Conroy C., Johnson B. D., 2017, *ApJ*, **840**, 44

- Calzetti D., Kinney A. L., Storchi-Bergmann T., 1994, *ApJ*, **429**, 582
- Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, *ApJ*, **533**, 682
- Camps P., Baes M., 2015, *Astronomy and Computing*, **9**, 20
- Camps P., Baes M., 2020, *Astronomy and Computing*, **31**, 100381
- Camps P., Trayford J. W., Baes M., Theuns T., Schaller M., Schaye J., 2016, *MNRAS*, **462**, 1057
- Capak P. L., et al., 2015, *Nature*, **522**, 455
- Carniani S., et al., 2018, *MNRAS*, **478**, 1170
- Casey C. M., 2012, *MNRAS*, **425**, 3094
- Casey C. M., et al., 2014, *ApJ*, **796**, 95
- Casey C. M., et al., 2018, *ApJ*, **862**, 77
- Cen R., Kimm T., 2014, *The Astrophysical Journal*, **782**, 32
- Ceverino D., Klypin A., Klimek E. S., Trujillo-Gomez S., Churchill C. W., Primack J., Dekel A., 2014, *MNRAS*, **442**, 1545
- Ceverino D., Glover S. C. O., Klessen R. S., 2017, *MNRAS*, **470**, 2791
- Chabrier G., 2003, *PASP*, **115**, 763
- Chan T. K., Theuns T., Bower R., Frenk C., 2021, *MNRAS*, **505**, 5784
- Charlot S., Fall S. M., 2000, *ApJ*, **539**, 718
- Chiang Y.-K., Overzier R., Gebhardt K., 2013, *ApJ*, **779**, 127
- Chiang Y.-K., Overzier R. A., Gebhardt K., Henriques B., 2017, *ApJL*, **844**, L23
- Ciesla L., et al., 2014, *A&A*, **565**, A128
- Clay S. J., Thomas P. A., Wilkins S. M., Henriques B. M. B., 2015, *MNRAS*, **451**, 2692
- Clemens M. S., et al., 2013, *MNRAS*, **433**, 695
- Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, *MNRAS*, **319**, 168
- Conroy C., 2013, *ARA&A*, **51**, 393

- Conroy C., Gunn J. E., White M., 2009, *ApJ*, **699**, 486
- Correa C. A., Schaye J., Clauwens B., Bower R. G., Crain R. A., Schaller M., Theuns T., Thob A. C. R., 2017, *MNRAS*, **472**, L45
- Cortzen I., et al., 2020, *A&A*, **634**, L14
- Crain R. A., et al., 2009, *MNRAS*, **399**, 1773
- Crain R. A., et al., 2015, *MNRAS*, **450**, 1937
- Crain R. A., et al., 2017, *MNRAS*, **464**, 4204
- Croton D. J., et al., 2006, *MNRAS*, **365**, 11
- Croton D. J., et al., 2016, *ApJS*, **222**, 22
- Cullen L., Dehnen W., 2010, *MNRAS*, **408**, 669
- Dale D. A., et al., 2012, *ApJ*, **745**, 95
- Dalgarno A., McCray R. A., 1972, *ARA&A*, **10**, 375
- Dalla Vecchia C., Schaye J., 2012, *MNRAS*, **426**, 140
- Davé R., Thompson R. J., Hopkins P. F., 2016, *MNRAS*, **462**, 3265
- Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, *MNRAS*, **486**, 2827
- Davies J. I., et al., 2017, *PASP*, **129**, 044102
- Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, *ApJ*, **292**, 371
- Dayal P., Ferrara A., 2018, *Phys. Rep.*, **780**, 1
- Dayal P., et al., 2020, *MNRAS*, **495**, 3065
- De Barros S., Oesch P. A., Labbé I., Stefanon M., González V., Smit R., Bouwens R. J., Illingworth G. D., 2019, *MNRAS*, **489**, 2355
- De Boer K. S., Jura M. A., Shull J. M., 1987, *Diffuse and Dark Clouds in the Interstellar Medium*. Springer Netherlands, Dordrecht, pp 485–515
- De Cia A., Ledoux C., Mattsson L., Petitjean P., Srianand R., Gavignaud I., Jenkins E. B., 2016, *A&A*, **596**, A97

- De Looze I., Baes M., Bendo G. J., Cortese L., Fritz J., 2011, *MNRAS*, 416, 2712
- De Looze I., et al., 2014, *A&A*, 568, A62
- De Lucia G., Kauffmann G., White S. D. M., 2004, *MNRAS*, 349, 1101
- De Lucia G., Tornatore L., Frenk C. S., Helmi A., Navarro J. F., White S. D. M., 2014, *MNRAS*, 445, 970
- De Vis, P. et al., 2019, *A&A*, 623, A5
- Decarli R., et al., 2019, *ApJ*, 882, 138
- Di Matteo T., Croft R. A. C., Feng Y., Waters D., Wilkins S., 2017, *MNRAS*, 467, 4243
- Diemand J., Kuhlen M., Madau P., Zemp M., Moore B., Potter D., Stadel J., 2008, *Nature*, 454, 735
- Dolag K., Borgani S., Murante G., Springel V., 2009, *MNRAS*, 399, 497
- Draine B., Li A., 2007, *ApJ*, 657, 810
- Draine B. T., et al., 2014, *ApJ*, 780, 172
- Dubois Y., et al., 2014, *MNRAS*, 444, 1453
- Dullemond C. P., Juhasz A., Pohl A., Sereshti F., Shetty R., Peters T., Commercon B., Flock M., 2012, RADMC-3D: A multi-purpose radiative transfer tool (ascl:1202.015)
- Duncan K., et al., 2014, *MNRAS*, 444, 2960
- Dunlop J. S., McLure R. J., Robertson B. E., Ellis R. S., Stark D. P., Cirasuolo M., de Ravel L., 2012, *MNRAS*, 420, 901
- Dunlop J. S., et al., 2017, *MNRAS*, 466, 861
- Dunne L., Eales S. A., Edmunds M. G., 2003, *MNRAS*, 341, 589
- Dunne L., et al., 2011, *MNRAS*, 417, 1510
- Durier F., Dalla Vecchia C., 2012, *MNRAS*, 419, 465
- Dwek E., 1998, *ApJ*, 501, 643
- Dwek E., Staguhn J., Arendt R. G., Kovacks A., Su T., Benford D. J., 2014, *ApJL*, 788, L30
- Eales S., et al., 2009, *ApJ*, 707, 1779

- Endsley R., Stark D. P., Chevallard J., Charlot S., 2021, *MNRAS*, 500, 5229
- Faisst A. L., et al., 2016, *ApJ*, 822, 29
- Faisst A. L., et al., 2020a, *ApJS*, 247, 61
- Faisst A. L., Fudamoto Y., Oesch P. A., Scoville N., Riechers D. A., Pavesi R., Capak P., 2020b, *MNRAS*, 498, 4192
- Feng Y., Di-Matteo T., Croft R. A., Bird S., Battaglia N., Wilkins S., 2016, *MNRAS*, 455, 2778
- Ferland G. J., et al., 2017, *Rev. Mex. Astron. Astrofis.*, 53, 385
- Ferrara A., Hirashita H., Ouchi M., Fujimoto S., 2017, *MNRAS*, 471, 5018
- Ferrarotti A. S., Gail H.-P., 2006, *A&A*, 447, 553
- Finkelstein S. L., et al., 2012, *ApJ*, 756, 164
- Finkelstein S. L., et al., 2015, *ApJ*, 810, 71
- Finlator K., Keating L., Oppenheimer B. D., Davé R., Zackrisson E., 2018, *MNRAS*, 480, 2628
- Fioc M., Rocca-Volmerange B., 1997, *A&A*, 500, 507
- Fitzpatrick E. L., Massa D., 2007, *ApJ*, 663, 320
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306
- Franco M., et al., 2018, *A&A*, 620, A152
- Fu J., et al., 2013, *MNRAS*, 434, 1531
- Fudamoto Y., et al., 2020, *A&A*, 643, A4
- Furlong M., et al., 2015, *MNRAS*, 450, 4486
- Furlong M., et al., 2017, *MNRAS*, 465, 722
- Galliano, F. et al., 2011, *A&A*, 536, A88
- Galliano F., Galametz M., Jones A. P., 2018, *ARA&A*, 56, 673
- Genel S., et al., 2014, *MNRAS*, 445, 175
- Giallongo E., et al., 2015, *A&A*, 578, A83

- Gjergo E., Granato G. L., Murante G., Ragone-Figueroa C., Tornatore L., Borgani S., 2018, *MNRAS*, 479, 2588
- Glikman E., Djorgovski S. G., Stern D., Dey A., Jannuzi B. T., Lee K.-S., 2011, *ApJ*, 728, L26
- Gnedin N. Y., 2000, *ApJ*, 542, 535
- Gnedin N. Y., 2014, *ApJ*, 793, 29
- González-López J., et al., 2019, *ApJ*, 882, 139
- Gonzalez-Perez V., Lacey C. G., Baugh C. M., Lagos C. D. P., Helly J., Campbell D. J. R., Mitchell P. D., 2014, *MNRAS*, 439, 264
- Gonzalez V., Labbe I., Bouwens R., Illingworth G., Franx M., Kriek M., 2011, *ApJ*, 735, L34
- Goodman J., Weare J., 2010, *Communications in Applied Mathematics and Computational Science*, 5, 65
- Grand R. J. J., et al., 2017, *MNRAS*, 467, 179
- Graziani L., Schneider R., Ginolfi M., Hunt L. K., Maio U., Glatzle M., Ciardi B., 2020, *MNRAS*, 494, 1071
- Greve T. R., et al., 2005, *MNRAS*, 359, 1165
- Groves B. A., Dopita M. A., Sutherland R. S., 2004a, *ApJS*, 153, 9
- Groves B. A., Dopita M. A., Sutherland R. S., 2004b, *ApJS*, 153, 75
- Groves B., Dopita M. A., Sutherland R. S., Kewley L. J., Fischera J., Leitherer C., Brandl B., van Breugel W., 2008, *The Astrophysical Journal Supplement Series*, 176, 438
- Gruppioni C., et al., 2013, *MNRAS*, 432, 23
- Gruppioni C., et al., 2020, *A&A*, 643, A8
- Guo Q., et al., 2011, *MNRAS*, 413, 101
- Guth A. H., 1981, *Phys. Rev. D*, 23, 347
- Gutiérrez C. M., López-Corredoira M., 2014, *A&A*, 571, A66
- Gutkin J., Charlot S., Bruzual G., 2016, *MNRAS*, 462, 1757

- Haardt F., Madau P., 2001, in Neumann D. M., Tran J. T. V., eds, Clusters of Galaxies and the High Redshift Universe Observed in X-rays. p. 64 ([arXiv:astro-ph/0106018](#))
- Harikane Y., et al., 2020, [ApJ](#), **896**, 93
- Harris C. R., et al., 2020, [Nature](#), 585, 357–362
- Hashimoto T., et al., 2018, [Nature](#), **557**, 392
- Hashimoto T., et al., 2019, [PASJ](#), **71**, 71
- Hayward C. C., Kereš D., Jonsson P., Narayanan D., Cox T. J., Hernquist L., 2011, [ApJ](#), **743**, 159
- Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, [MNRAS](#), **451**, 2663
- Henriques B. M. B., Yates R. M., Fu J., Guo Q., Kauffmann G., Srisawat C., Thomas P. A., White S. D. M., 2020, [MNRAS](#), **491**, 5795
- Hirashita H., Murga M. S., 2020, [MNRAS](#), **492**, 3779
- Hirashita H., Aoyama S., Hou K.-C., Shimizu I., Nagamine K., 2018, [MNRAS](#), **478**, 4905
- Hirschmann M., Charlot S., Feltre A., Naab T., Choi E., Ostriker J. P., Somerville R. S., 2017, [MNRAS](#), **472**, 2468
- Hirschmann M., Charlot S., Feltre A., Naab T., Somerville R. S., Choi E., 2019, [MNRAS](#), **487**, 333
- Hockney R. W., Eastwood J. W., 1988, Computer simulation using particles
- Hodge J. A., da Cunha E., 2020, [Royal Society Open Science](#), **7**, 200556
- Hodge J. A., et al., 2013, [ApJ](#), **768**, 91
- Hopkins P. F., 2013, [MNRAS](#), **428**, 2840
- Hopkins P. F., 2015, [MNRAS](#), **450**, 53
- Hopkins P. F., et al., 2018, [MNRAS](#), **480**, 800
- Hou K.-C., Hirashita H., Nagamine K., Aoyama S., Shimizu I., 2017, [MNRAS](#), **469**, 870
- Hou K.-C., Aoyama S., Hirashita H., Nagamine K., Shimizu I., 2019, [MNRAS](#), **485**, 1727
- Hu C.-Y., Zhukovska S., Somerville R. S., Naab T., 2019, [MNRAS](#), **487**, 3252
- Hunter J. D., 2007, [Computing in Science & Engineering](#), **9**, 90

- Hutchison T. A., et al., 2019, *ApJ*, 879, 70
- Hutter A., Dayal P., Yepes G., Gottlöber S., Legrand L., Ucci G., 2021, *MNRAS*, 503, 3698
- Ilbert O., et al., 2013, *A&A*, 556, A55
- Inoue A. K., 2003, *PASJ*, 55, 901
- Inoue A. K., et al., 2016, *Science*, 352, 1559
- Irvine W. M., Goldsmith P. F., Hjalmarson Å., 1987, in Hollenbach D. J., Thronson H. A., eds, *Interstellar Processes*. Springer Netherlands, Dordrecht, pp 560–609
- Ito K., et al., 2020, *ApJ*, 899, 5
- Jaacks J., Finkelstein S. L., Bromm V., 2019, *MNRAS*, 488, 2202
- Jenkins E. B., 2009, *ApJ*, 700, 1299
- Jenkins A., 2010, *MNRAS*, 403, 1859
- Jenkins A., 2013, *MNRAS*, 434, 2094
- Jin S., et al., 2019, *ApJ*, 887, 144
- Jones A. P., 2000, *Journal of Geophysical Research: Space Physics*, 105, 10257
- Jonsson P., 2006, *MNRAS*, 372, 2
- Jonsson P., Groves B. A., Cox T. J., 2010, *MNRAS*, 403, 17
- Kannan R., Vogelsberger M., Marinacci F., McKinnon R., Pakmor R., Springel V., 2019, *MNRAS*, 485, 117
- Katsianis A., et al., 2017, *MNRAS*, 472, 919
- Katz H., Kimm T., Sijacki D., Haehnelt M. G., 2017, *MNRAS*, 468, 4831
- Katz H., Kimm T., Haehnelt M., Sijacki D., Rosdahl J., Blaizot J., 2018, *MNRAS*, 478, 4986
- Kauffmann G., Charlot S., White S. D. M., 1996, *MNRAS*, 283, L117
- Kauffmann G., Colberg J. M., Diaferio A., White S. D. M., 1999, *MNRAS*, 303, 188
- Kawamata R., Ishigaki M., Shimasaku K., Oguri M., Ouchi M., Tanigawa S., 2018, *ApJ*, 855, 4
- Kelvin L. S., et al., 2014, *MNRAS*, 444, 1647

- Kennicutt Jr R. C., Evans II N. J., 2012, *ARA&A*, 50, 531
- Kennicutt R. C., et al., 2011, *PASP*, 123, 1347
- Kewley L. J., Ellison S. L., 2008, *ApJ*, 681, 1183
- Kewley L. J., Nicholls D. C., Sutherland R. S., 2019, *ARA&A*, 57, 511
- Khandai N., Di Matteo T., Croft R., Wilkins S., Feng Y., Tucker E., DeGraf C., Liu M.-S., 2015, *MNRAS*, 450, 1349
- Khusanova Y., et al., 2020, arXiv e-prints, p. [arXiv:2007.08384](https://arxiv.org/abs/2007.08384)
- Klaassen P., et al., 2019, in Bulletin of the American Astronomical Society. p. 58 ([arXiv:1907.04756](https://arxiv.org/abs/1907.04756))
- Klessen R., 2019, Formation of the first stars. pp 67–97, [doi:10.1142/9789813227958_0004](https://doi.org/10.1142/9789813227958_0004)
- Knudsen K. K., Richard J., Kneib J.-P., Jauzac M., Clément B., Drouart G., Egami E., Lindroos L., 2016, *MNRAS*, 462, L6
- Knudsen K. K., Watson D., Frayer D., Christensen L., Gallazzi A., Michałowski M. J., Richard J., Zavala J., 2017, *MNRAS*, 466, 138
- Koprowski M. P., Dunlop J. S., Michałowski M. J., Coppin K. E. K., Geach J. E., McLure R. J., Scott D., van der Werf P. P., 2017, *MNRAS*, 471, 4155
- Koprowski M. P., et al., 2018, *MNRAS*, 479, 4355
- Kroupa P., 2001, *MNRAS*, 322, 231
- Kroupa P., 2002, *Science*, 295, 82
- Labbé I., et al., 2010, *ApJ*, 708, L26
- Lacey C. G., Baugh C. M., Frenk C. S., Benson a. J., 2011, *MNRAS*, 412, 1828
- Lacey C. G., et al., 2016, *MNRAS*, 462, 3854
- Lagache G., Cousin M., Chatzikos M., 2018, *A&A*, 609, A130
- Lagos C. d. P., et al., 2015, *MNRAS*, 452, 3815
- Lagos C. d. P., et al., 2019, *MNRAS*, 489, 4196
- Lakićević M., et al., 2015, *ApJ*, 799, 50

- Laporte N., et al., 2017, *The Astrophysical Journal*, 837, L21
- Le Brun A. M. C., McCarthy I. G., Schaye J., Ponman T. J., 2014, *Monthly Notices of the Royal Astronomical Society*, 441, 1270
- Le Fèvre O., et al., 2020, *A&A*, 643, A1
- Leitherer C., et al., 1999, *ApJS*, 123, 3
- Leung T. K. D., Olsen K. P., Somerville R. S., Davé R., Greve T. R., Hayward C. C., Narayanan D., Popping G., 2020, *ApJ*, 905, 102
- Li Q., Narayanan D., Davé R., 2019, *MNRAS*, 490, 1425
- Liang L., et al., 2019, *MNRAS*, 489, 1397
- Liang L., Feldmann R., Hayward C. C., Narayanan D., Çatmabacak O., Kereš D., Faucher-Giguère C.-A., Hopkins P. F., 2021, *MNRAS*, 502, 3210
- Liddle A. R., 2007, *MNRAS*, 377, L74
- Linde A. D., 1982, *Physics Letters B*, 108, 389
- Lovell C. C., Thomas P. A., Wilkins S. M., 2018, *MNRAS*, 474, 4612
- Lovell C. C., Vijayan A. P., Thomas P. A., Wilkins S. M., Barnes D. J., Irodotou D., Roper W., 2021a, *MNRAS*, 500, 2127
- Lovell C. C., Geach J. E., Davé R., Narayanan D., Li Q., 2021b, *MNRAS*, 502, 772
- Ma X., et al., 2018, *MNRAS*, 478, 1694
- Ma X., et al., 2019, *MNRAS*, 487, 1844
- Ma X., Quataert E., Wetzel A., Hopkins P. F., Faucher-Giguère C.-A., Kereš D., 2020, *MNRAS*,
- Madden S. C., et al., 2013, *PASP*, 125, 600
- Magdis G. E., et al., 2012, *ApJ*, 760, 6
- Magnelli B., et al., 2014, *A&A*, 561, A86
- Mancini M., Schneider R., Graziani L., Valiante R., Dayal P., Maio U., Ciardi B., Hunt L. K., 2015, *MNRAS*, 451, L70

- Mancini M., Schneider R., Graziani L., Valiante R., Dayal P., Maio U., Ciardi B., 2016, *MNRAS*, 462, 3130
- Maraston C., 2005, *MNRAS*, 362, 799
- Marigo P., 2001, *A&A*, 370, 194
- Marinacci F., et al., 2018, *MNRAS*, 480, 5113
- Marrone D. P., et al., 2018, *Nature*, 553, 51
- Martindale H., Thomas P. A., Henriques B. M., Loveday J., 2017, *MNRAS*, 472, 1981
- Mason C. A., Trenti M., Treu T., 2015, *ApJ*, 813, 21
- Matteo T. D., Khandai N., DeGraf C., Feng Y., Croft R. A. C., Lopez J., Springel V., 2012, *ApJ*, 745, L29
- Mattsson L., 2015, preprint, ([arXiv:1505.04758](https://arxiv.org/abs/1505.04758))
- McAlpine S., et al., 2016, *A&C*, 15, 72
- McAlpine S., et al., 2019, *MNRAS*, 488, 2440
- McKee C. F., 1989, *ApJ*, 345, 782
- McKee C. F., Krumholz M. R., 2009, *ApJ*, 709, 308
- McKinnon R., Torrey P., Vogelsberger M., 2016, *MNRAS*, 457, 3775
- McKinnon R., Torrey P., Vogelsberger M., Hayward C. C., Marinacci F., 2017, *MNRAS*, 468, 1505
- McKinnon R., Vogelsberger M., Torrey P., Marinacci F., Kannan R., 2018, *MNRAS*, 478, 2851
- McLeod D. J., McLure R. J., Dunlop J. S., Robertson B. E., Ellis R. S., Targett T. A., 2015, *MNRAS*, 450, 3032
- Meiksin A., Madau P., 1993, *ApJ*, 412, 34
- Merlin E., et al., 2019, *MNRAS*, 490, 3309
- Meurer G. R., Heckman T. M., Calzetti D., 1999, *ApJ*, 521, 64
- Michałowski M. J., 2015, *A&A*, 577, A80
- Milgrom M., 1983, *ApJ*, 270, 365

- Moffett A. J., et al., 2016, *MNRAS*, 457, 1308
- Monaghan J. J., 1992, *ARA&A*, 30, 543
- Morgan H. L., Edmunds M. G., 2003, *MNRAS*, 343, 427
- Moriwaki K., et al., 2018, *MNRAS*, 481, L84
- Mortlock D. J., et al., 2011, *Nature*, 474, 616
- Motte F., et al., 2018, *Nature Astronomy*, 2, 478
- Mukhanov V. F., Chibisov G. V., 1981, Soviet Journal of Experimental and Theoretical Physics Letters, 33, 532
- Naiman J. P., et al., 2018, *MNRAS*, 477, 1206
- Narayanan D., Davé R., Johnson B. D., Thompson R., Conroy C., Geach J., 2018, *MNRAS*, 474, 1718
- Narayanan D., et al., 2021, *ApJS*, 252, 12
- Nelson D., et al., 2018, *MNRAS*, 475, 624
- Newville M., Stensitzki T., Allen D. B., Ingargiola A., 2014, LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python, doi:10.5281/zenodo.11813, <https://doi.org/10.5281/zenodo.11813>
- Nozawa T., Maeda K., Kozasa T., Tanaka M., Nomoto K., Umeda H., 2011, *ApJ*, 736, 45
- O’Shea B. W., Wise J. H., Xu H., Norman M. L., 2015, *ApJ*, 807, L12
- Ocvirk P., et al., 2016, *MNRAS*, 463, 1462
- Ocvirk P., et al., 2020, *MNRAS*, 496, 4087
- Oesch P. A., et al., 2016, *ApJ*, 819, 129
- Oesch P. A., Bouwens R. J., Illingworth G. D., Labbé I., Stefanon M., 2018, *ApJ*, 855, 105
- Okamoto T., Gao L., Theuns T., 2008, *MNRAS*, 390, 920
- Olsen K., Greve T. R., Narayanan D., Thompson R., Davé R., Niebla Rios L., Stawinski S., 2017, *ApJ*, 846, 105
- Ono Y., et al., 2018, *PASJ*, 70, S10

- Ota K., et al., 2014, *ApJ*, 792, 34
- Ouchi M., et al., 2013, *ApJ*, 778, 102
- Pallottini A., et al., 2019, *MNRAS*, 487, 1689
- Peebles P. J., Ratra B., 2003, *Reviews of Modern Physics*, 75, 559
- Peek J. E. G., Ménard B., Corrales L., 2015, *ApJ*, 813, 7
- Pei Y. C., 1992, *ApJ*, 395, 130
- Pettini M., Kellogg M., Steidel C. C., Dickinson M., Adelberger K. L., Giavalisco M., 1998, *The Astrophysical Journal*, 508, 539
- Pike S. R., Kay S. T., Newton R. D. A., Thomas P. A., Jenkins A., 2014, *MNRAS*, 445, 1774
- Pillepich A., et al., 2018, *MNRAS*, 475, 648
- Planck Collaboration et al., 2014, *A&A*, 571, A1
- Planelles S., Borgani S., Fabjan D., Killevar M., Murante G., Granato G. L., Ragone-Figueroa C., Dolag K., 2014, *MNRAS*, 438, 195
- Poole G. B., Angel P. W., Mutch S. J., Power C., Duffy A. R., Geil P. M., Mesinger A., Wyithe S. B., 2016, *MNRAS*, 459, 3025
- Popping G., van Kampen E., Decarli R., Spaans M., Somerville R. S., Trager S. C., 2016, *MNRAS*, 461, 93
- Popping G., Somerville R. S., Galametz M., 2017a, *MNRAS*, 471, 3152
- Popping G., Puglisi A., Norman C. A., 2017b, *MNRAS*, 472, 2315
- Popping G., Narayanan D., Somerville R. S., Faisst A. L., Krumholz M. R., 2019, *MNRAS*, 482, 4906
- Portinari L., Chiosi C., Bressan A., 1998, *A&A*, 334, 505
- Price D. J., 2008, *Journal of Computational Physics*, 227, 10040
- Reddy N. A., et al., 2015, *ApJ*, 806, 259
- Rémy-Ruyer A., et al., 2014, *A&A*, 563, A31
- Rémy-Ruyer A., et al., 2015, *A&A*, 582, A121
- Roberts-Borsani G. W., et al., 2016, *ApJ*, 823, 143

- Robertson B. E., 2010, *ApJ*, 713, 1266
- Robertson B. E., Ellis R. S., Dunlop J. S., McLure R. J., Stark D. P., 2010, *Nature*, 468, 49
- Robertson B. E., et al., 2013, *ApJ*, 768, 71
- Robertson B. E., Ellis R. S., Furlanetto S. R., Dunlop J. S., 2015, *ApJ*, 802, L19
- Rodrigues L. F. S., Vernon I., Bower R. G., 2017, *MNRAS*, 466, 2418
- Rohatgi A., 2020, Webplotdigitizer: Version 4.4, <https://automeris.io/WebPlotDigitizer>
- Rosas-Guevara Y. M., et al., 2015, *MNRAS*, 454, 1038
- Rosdahl J., Teyssier R., 2015, *MNRAS*, 449, 4380
- Rosdahl J., et al., 2018, *MNRAS*, 479, 994
- Safarzadeh M., Hayward C. C., Ferguson H. C., 2017, *The Astrophysical Journal*, 840, 15
- Salim S., Lee J. C., 2012, *ApJ*, 758, 134
- Salim S., Narayanan D., 2020, *ARA&A*, 58, 529
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Sanders D. B., Mirabel I. F., 1996, *ARA&A*, 34, 749
- Santini P., et al., 2014, *A&A*, 562, A30
- Savage B. D., Sembach K. R., 1996, *ApJ*, 470, 893
- Sawala T., et al., 2016, *MNRAS*, 457, 1931
- Sawala T., McAlpine S., Jasche J., Lavaux G., Jenkins A., Johansson P. H., Frenk C. S., 2021, arXiv e-prints, p. [arXiv:2103.12073](https://arxiv.org/abs/2103.12073)
- Schaerer D., Boone F., Zamojski M., Staguhn J., Dessauges-Zavadsky M., Finkelstein S., Combes F., 2015, *A&A*, 574, A19
- Schaller M., Dalla Vecchia C., Schaye J., Bower R. G., Theuns T., Crain R. A., Furlong M., McCarthy I. G., 2015, *MNRAS*, 454, 2277
- Schaye J., 2004, *ApJ*, 609, 667
- Schaye J., Dalla Vecchia C., 2008, *MNRAS*, 383, 1210

- Schaye J., et al., 2010, *MNRAS*, 402, 1536
- Schaye J., et al., 2015, *MNRAS*, 446, 521
- Schechter P., 1976, *ApJ*, 203, 297
- Schneider F. R. N., et al., 2018, *Science*, 359, 69
- Schouws S., et al., 2021, arXiv e-prints, p. [arXiv:2105.12133](#)
- Schreiber C., Elbaz D., Pannella M., Ciesla L., Wang T., Franco M., 2018, *A&A*, 609, A30
- Schulz S., Popping G., Pillepich A., Nelson D., Vogelsberger M., Marinacci F., Hernquist L., 2020, *MNRAS*, 497, 4773
- Schwarz G., 1978, *Annals of Statistics*, 6, 461
- Scott K. S., et al., 2011, *ApJ*, 733, 29
- Scoville N., et al., 2016, *The Astrophysical Journal*, 820, 83
- Searle L., Sargent W. L. W., Bagnuolo W. G., 1973, *ApJ*, 179, 427
- Shen X., et al., 2020, *MNRAS*, 495, 4747
- Shen X., Vogelsberger M., Nelson D., Tacchella S., Hernquist L., Springel V., Marinacci F., Torrey P., 2021, arXiv e-prints, p. [arXiv:2104.12788](#)
- Shibuya T., Miura N., Iwadate K., Fujimoto S., Harikane Y., Toba Y., Umayahara T., Ito Y., 2021, arXiv e-prints, p. [arXiv:2106.03728](#)
- Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, *MNRAS*, 452, 575
- Simpson J. M., et al., 2014, *ApJ*, 788, 125
- Smit R., Bouwens R. J., Franx M., Illingworth G. D., Labbé I., Oesch P. A., Dokkum P. G. v., 2012, *ApJ*, 756, 14
- Smit R., et al., 2018, *Nature*, 553, 178
- Somerville R. S., Davé R., 2015, *ARA&A*, 53, 51
- Somerville R. S., Primack J. R., 1999, *MNRAS*, 310, 1087
- Somerville R. S., Primack J. R., Faber S. M., 2001, *MNRAS*, 320, 504

- Somerville R. S., Popping G., Trager S. C., 2015, *MNRAS*, 453, 4337
- Sommovigo L., Ferrara A., Pallottini A., Carniani S., Gallerani S., Decataldo D., 2020, *MNRAS*, 497, 956
- Song M., et al., 2016, *ApJ*, 825, 5
- Springel V., 2010a, *ARA&A*, 48, 391
- Springel V., 2010b, *MNRAS*, 401, 791
- Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, *MNRAS*, 328, 726
- Springel V., et al., 2005a, *Nature*
- Springel V., Di Matteo T., Hernquist L., 2005b, *MNRAS*, 361, 776
- Springel V., et al., 2008, *MNRAS*, 391, 1685
- Springel V., et al., 2018, *MNRAS*, 475, 676
- Stacey G. J., Geis N., Genzel R., Lugten J. B., Poglitsch A., Sternberg A., Townes C. H., 1991, *ApJ*, 373, 423
- Stalevski M., 2012, *Bulgarian Astronomical Journal*, 18, 3
- Stalevski M., Ricci C., Ueda Y., Lira P., Fritz J., Baes M., 2016, *MNRAS*, 458, 2288
- Stanway E. R., 2017,] 10.1017/S1743921317002927, 329, 305
- Stanway E. R., Eldridge J. J., 2018, *MNRAS*, 479, 75
- Stanway E. R., McMahon R. G., Bunker A. J., 2005, *MNRAS*, 359, 1184
- Stark D. P., et al., 2015, *MNRAS*, 450, 1846
- Stark D. P., et al., 2017, *MNRAS*, 464, 469
- Stefanon M., Bouwens R. J., Labbé I., Muzzin A., Marchesini D., Oesch P., Gonzalez V., 2017, *ApJ*, 843, 36
- Stefanon M., et al., 2019, *ApJ*, 883, 99
- Steigman G., 2007, *Annual Review of Nuclear and Particle Science*, 57, 463
- Strandet M. L., et al., 2016, *The Astrophysical Journal*, 822, 80

- Swinbank A. M., et al., 2012, *MNRAS*, 427, 1066
- Tacconi L. J., et al., 2006, *ApJ*, 640, 228
- Takeuchi T. T., Yuan F.-T., Ikeyama A., Murata K. L., Inoue A. K., 2012, *ApJ*, 755, 144
- Tamura Y., et al., 2019, *The Astrophysical Journal*, 874, 27
- Teyssier R., 2002, *A&A*, 385, 337
- Thielemann F.-K., et al., 2003, *Nuclear Physics A*, 718, 139
- Tinsley B. M., 1972, *A&A*, 20, 383
- Tinsley B. M., 1980, *A&A*, 89, 246
- Tomczak A. R., et al., 2016, *ApJ*, 817, 118
- Trayford J. W., et al., 2015, *MNRAS*, 452, 2879
- Trayford J. W., et al., 2017, *MNRAS*, 470, 771
- Tremmel M., Karcher M., Governato F., Volonteri M., Quinn T. R., Pontzen A., Anderson L., Bellovary J., 2017, *MNRAS*, 470, 1121
- Triani D. P., Sinha M., Croton D. J., Pacifici C., Dwek E., 2020, *MNRAS*, 493, 2490
- Troncoso P., et al., 2014, *A&A*, 563, A58
- Trčka A., et al., 2020, *MNRAS*, 494, 2823
- Tsai J. C., Mathews W. G., 1995, *ApJ*, 448, 84
- Vallini L., Gallerani S., Ferrara A., Pallottini A., Yue B., 2015, *ApJ*, 813, 36
- Venemans B. P., et al., 2012, *ApJ*, 751, L25
- Vijayan A. P., Clay S. J., Thomas P. A., Yates R. M., Wilkins S. M., Henriques B. M., 2019, *MNRAS*, 489, 4072
- Vijayan A. P., Lovell C. C., Wilkins S. M., Thomas P. A., Barnes D. J., Irodotou D., Kuusisto J., Roper W. J., 2021, *MNRAS*, 501, 3289
- Virtanen P., et al., 2020, *Nature Methods*, 17, 261
- Vlahakis C., Dunne L., Eales S., 2005, *MNRAS*, 364, 1253

- Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, *MNRAS*, **436**, 3031
- Vogelsberger M., et al., 2014a, *MNRAS*, **444**, 1518
- Vogelsberger M., et al., 2014b, *Nature*, **509**, 177
- Vogelsberger M., Marinacci F., Torrey P., Puchwein E., 2020a, *Nature Reviews Physics*, **2**, 42
- Vogelsberger M., et al., 2020b, *MNRAS*, **492**, 5167
- Wagoner R. V., Fowler W. A., Hoyle F., 1967, *ApJ*, **148**, 3
- Walcher J., Groves B., Budavári T., Dale D., 2011, *Ap&SS*, **331**, 1
- Walter F., et al., 2016, *ApJ*, **833**, 67
- Wang W., et al., 2018, *The Astrophysical Journal*, **869**, 161
- Wang L., Pearson W. J., Cowley W., Trayford J. W., Béthermin M., Gruppioni C., Hurley P., Michałowski M. J., 2019, *A&A*, **624**, A98
- Waters D., Di Matteo T., Feng Y., Wilkins S. M., Croft R. A. C., 2016, *MNRAS*, **463**, 3520
- Watson D., Christensen L., Knudsen K. K., Richard J., Gallazzi A., Michałowski M. J., 2015, *NATURE*, **519**, 327
- Weingartner J. C., Draine B. T., 2001, *ApJ*, **548**, 296
- Wetzel A. R., Hopkins P. F., Kim J.-h., Faucher-Giguère C.-A., Kereš D., Quataert E., 2016, *ApJ*, **827**, L23
- White S. D. M., Frenk C. S., 1991, *ApJ*, **379**, 52
- Wiersma R. P. C., Schaye J., Smith B. D., 2009a, *MNRAS*, **393**, 99
- Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009b, *MNRAS*, **399**, 574
- Wilkins S. M., Bunker A. J., Ellis R. S., Stark D., Stanway E. R., Chiu K., Lorenzoni S., Jarvis M. J., 2010, *MNRAS*, **403**, 938
- Wilkins S. M., Bunker A. J., Lorenzoni S., Caruana J., 2011a, *MNRAS*, **411**, 23
- Wilkins S. M., Bunker A. J., Stanway E., Lorenzoni S., Caruana J., 2011b, *MNRAS*, **417**, 717
- Wilkins S. M., Gonzalez-Perez V., Lacey C. G., Baugh C. M., 2012, *MNRAS*, **424**, 1522

- Wilkins S. M., Bunker A., Coulton W., Croft R., Matteo T. D., Khandai N., Feng Y., 2013, *MNRAS*, **430**, 2885
- Wilkins S. M., Bouwens R. J., Oesch P. A., Labbé I., Sargent M., Caruana J., Wardlow J., Clay S., 2016a, *MNRAS*, **455**, 659
- Wilkins S. M., Feng Y., Di Matteo T., Croft R., Stanway E. R., Bouwens R. J., Thomas P., 2016b, *MNRAS*, **458**, L6
- Wilkins S. M., Feng Y., Di-Matteo T., Croft R., Stanway E. R., Bunker A., Waters D., Lovell C., 2016c, *MNRAS*, **460**, 3170
- Wilkins S. M., Feng Y., Di Matteo T., Croft R., Lovell C. C., Waters D., 2017, *MNRAS*, **469**, 2517
- Wilkins S. M., Feng Y., Di Matteo T., Croft R., Lovell C. C., Thomas P., 2018, *MNRAS*, **473**, 5363
- Wilkins S. M., et al., 2020, *MNRAS*, **493**, 6079
- Wiseman P., Schady P., Bolmer J., Krühler T., Yates R. M., Greiner J., Fynbo J. P. U., 2017, *A&A*, **599**, A24
- Wu X., Kannan R., Marinacci F., Vogelsberger M., Hernquist L., 2019, *MNRAS*, **488**, 419
- Wu X., Davé R., Tacchella S., Lotz J., 2020, *MNRAS*, **494**, 5636
- Yates R. M., Henriques B., Thomas P. A., Kauffmann G., Johansson J., White S. D. M., 2013, *MNRAS*, **435**, 3500
- Yates R. M., Thomas P. A., Henriques B. M. B., 2017, *MNRAS*, **464**, 3169
- Yoshida N., Hosokawa T., Omukai K., 2012, *Progress of Theoretical and Experimental Physics*, **2012**, 01A305
- Yung L. Y. A., Somerville R. S., Finkelstein S. L., Popping G., Davé R., 2019a, *MNRAS*, **483**, 2983
- Yung L. Y. A., Somerville R. S., Popping G., Finkelstein S. L., Ferguson H. C., Davé R., 2019b, *MNRAS*, **490**, 2855
- Zaroubi S., 2013, The Epoch of Reionization. p. 45, doi:10.1007/978-3-642-32362-1_2
- Zavala J. A., et al., 2021, *ApJ*, **909**, 165
- Zhang Z.-Y., Romano D., Ivison R. J., Papadopoulos P. P., Matteucci F., 2018, *Nature*, **558**, 260
- Zhukovska S., 2014, *A&A*, **562**, A76

Zhukovska S., Henning T., 2013, *A&A*, 555, A99

Zhukovska S., Gail H.-P., Tieloff M., 2008, *A&A*, 479, 453

Zubko V., Dwek E., Arendt R. G., 2004, *ApJS*, 152, 211

da Cunha E., Charlot S., Elbaz D., 2008, *MNRAS*, 388, 1595

da Cunha E., et al., 2013, *ApJ*, 766, 13

da Cunha E., et al., 2015, *ApJ*, 806, 110

van Dishoeck E. F., Blake G. A., 1998, *ARA&A*, 36, 317